PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 1/04

A2

(11) International Publication Number: WO 97/43301

(43) International Publication Date: 20 November 1997 (20.11.97)

(21) International Application Number: PCT/EP97/02215

(22) International Filing Date: 30 April 1997 (30.04.97)

(30) Priority Data:
60/017,545 10 May 1996 (10.05.96) US

(71) Applicant *(for all designated States except US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).

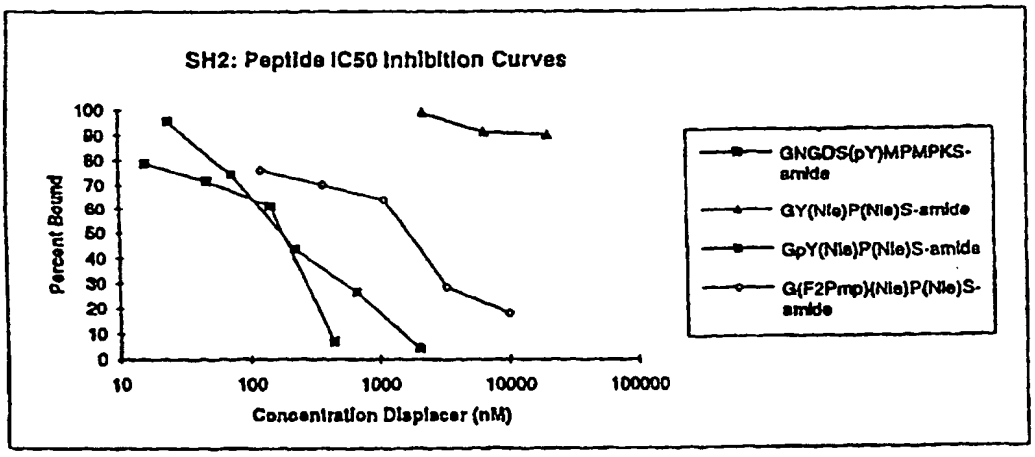
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WENNOGLE, Lawrence Paul [US/US]; 38 Inverrary Place, Annandale, NJ 08801 (US). KELLY, Michele, Ann [US/US]; 22 Van Ness Court, Maplewood, NJ 07040 (US). LIANG, Hongbin [CN/US]; Apartment A, 98 Gales Drive, New Providence, NJ 07974 (US). GOELLER, Christine [US/US]; 78 Westover Avenue, West Caldwell, NJ 07006 (US). THOMA, Hans, Mathis [CN/US]; 34 Ashland Road, Summit, NJ 07901 (US).

(74) Agent: ROTH, Bernhard, M.; Novartis AG, Patent- und Markenabteilung, Klybeckstrasse 141, CH-4002 Basel (CH).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: IDENTIFICATION OF MEMBERS OF COMBINATORIAL LIBRARIES BY MASS SPECTROMETRY




| | IC50 (nM) Ave. | SD |
|---|---|---|
| GNGDS(pY)MPMPKS-amide | 188 | 53 |
| GY(Nle)P(Nle)S-amide | >20000 | |
| GpY(Nle)P(Nle)S-amide | 202 | 24 |
| G(F2Pmp)(Nle)P(Nle)S-amide | 2881 | 2958 |

(57) Abstract

The present invention is drawn to methods for characterizing the members of a combinatorial library which bind to a domain of interest. The method utilizes affinity selection in combination with mass spectrometry to provide rapid and efficient screening. The method provides information on relative affinities and molecular weights of affinity-selected compounds. The methods find use in analyzing all types of combinatorial libraries.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | |
|---|---|
| AL | Albania |
| AM | Armenia |
| AT | Austria |
| AU | Australia |
| AZ | Azerbaijan |
| BA | Bosnia and Herzegovina |
| BB | Barbados |
| BE | Belgium |
| BF | Burkina Faso |
| BG | Bulgaria |
| BJ | Benin |
| BR | Brazil |
| BY | Belarus |
| CA | Canada |
| CF | Central African Republic |
| CG | Congo |
| CH | Switzerland |
| CI | Côte d'Ivoire |
| CM | Cameroon |
| CN | China |
| CU | Cuba |
| CZ | Czech Republic |
| DE | Germany |
| DK | Denmark |
| EE | Estonia |
| ES | Spain |
| FI | Finland |
| FR | France |
| GA | Gabon |
| GB | United Kingdom |
| GE | Georgia |
| GH | Ghana |
| GN | Guinea |
| GR | Greece |
| HU | Hungary |
| IE | Ireland |
| IL | Israel |
| IS | Iceland |
| IT | Italy |
| JP | Japan |
| KE | Kenya |
| KG | Kyrgyzstan |
| KP | Democratic People's Republic of Korea |
| KR | Republic of Korea |
| KZ | Kazakstan |
| LC | Saint Lucia |
| LI | Liechtenstein |
| LK | Sri Lanka |
| LR | Liberia |
| LS | Lesotho |
| LT | Lithuania |
| LU | Luxembourg |
| LV | Latvia |
| MC | Monaco |
| MD | Republic of Moldova |
| MG | Madagascar |
| MK | The former Yugoslav Republic of Macedonia |
| ML | Mali |
| MN | Mongolia |
| MR | Mauritania |
| MW | Malawi |
| MX | Mexico |
| NE | Niger |
| NL | Netherlands |
| NO | Norway |
| NZ | New Zealand |
| PL | Poland |
| PT | Portugal |
| RO | Romania |
| RU | Russian Federation |
| SD | Sudan |
| SE | Sweden |
| SG | Singapore |
| SI | Slovenia |
| SK | Slovakia |
| SN | Senegal |
| SZ | Swaziland |
| TD | Chad |
| TG | Togo |
| TJ | Tajikistan |
| TM | Turkmenistan |
| TR | Turkey |
| TT | Trinidad and Tobago |
| UA | Ukraine |
| UG | Uganda |
| US | United States of America |
| UZ | Uzbekistan |
| VN | Viet Nam |
| YU | Yugoslavia |
| ZW | Zimbabwe |

# IDENTIFICATION OF MEMBERS OF COMBINATORIAL LIBRARIES BY MASS SPECTROMETRY

## FIELD OF THE INVENTION

The invention relates to the characterization of combinatorial libraries.

## BACKGROUND OF THE INVENTION

The screening of chemicals and natural materials such as fermentation broths and plant extracts, has long been used to identify novel lead compounds for drug discovery. Assays range in complexity from simple binding reactions to elaborate physiological preparations.

Ligands discovered in this manner can be useful agents if they mimic or block natural ligands, or if they interfere with the naturally occurring interactions of the biological target. Such ligands also provide a starting point for the engineering of molecules with more desirable properties. The chance of finding valuable ligands increases with the number of compounds screened. Therefore, it is desirable to screen massive libraries of compounds.

Combinatorial chemistry provides a promising approach to the synthesis of large collections of diverse molecules. In combinatorial chemistry, vast libraries of molecules having different chemical compositions can be synthesized simultaneously. Such random synthetic peptide or nonpeptide libraries prove especially useful to the pharmaceutical industry for the rapid identification of novel compounds that bind to high-affinity acceptor molecules such as receptors, enzymes and antibodies.

Combinatorial methods entail a series of chemical steps with multiple choices of chemical reagents for each step. The complexity, or number of members in a combinatorial library, is related to the product of the number of reagent choices or building blocks for each step of the synthesis and can therefore be quite large.

To date, most methods to screen such combinatorial libraries include the use of acceptor molecules labeled with fluorescent or other reporter groups. In one method, combinatorial libraries are synthesized on resin beads. The beads are exposed to labeled acceptor

molecules. Those with bound acceptor are identified by visual inspection, physically removed, and the peptide is sequenced directly.

Various techniques for the evaluation of combinatorial libraries, have been developed where several organizational approaches have been implemented to keep track of individual components within a library. The iterative approach (Dooley, C.T., Chung, N.N., Schiller, P.W., & Houghten, R.A. (1993) *Proc Natl Acad Sci USA* 90, 10811-5; Dooley, C.T., & Houghten, R.A. (1993) *Life Sci* 52, 1509-17; Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., & Cuervo, J.H. (1991) *Nature* 354, 84-6; Pinilla, C., Appel, J.R., & Houghten, R.A. (1993) *Gene* 128, 71-6) systematically screens and isolates "hits" at each stage of the synthetic process (Furka, A., Sebestyen, F., Asgedom, M., & Dibo, G. (1988) in *14th Int. Congr. Biochem.* pp 47, Prague). Spatial addressing of library components using VLSIPS (Fodor, S.P., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., & Solas, D. (1991) *Science* 251, 767-73; Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P., & Gordon, E.M. (1994) *J Med Chem* 37, 1233-51) allows for selection and detection of individual library components according to their fixed presence in an array. The one-bead, one-peptide approach (Chen, J.K., Lane, W.S., Brauer, A.W., Tanaka, A., & Schreiber, S.L. (1993) *J Am Chem Soc* 115, 12591-12592; Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., & Knapp, R.J. (1991) *Nature* 354, 82-4; Lebl, M., Krchnak, V., Sepetov, N.F., Seligmann, B., Strop, P., Felder, S., & Lam, K.S. (1995) *Biopolymers* 37, 177-98; Salmon, S.E., Lam, K.S., Lebl, M., Kandola, A., Khattri, P.S., Wade, S., Patek, M., Kocis, P., Krchnak, V., Thorpe, D., & *et al.* (1993) *Proc Natl Acad Sci USA* 90, 11708-12) enables an entire library to be screened by microscopically isolating individual beads via tagged receptors. Encoded libraries are used to eliminate ambiguity, where coding strands are incorporated into the library through the use of various tags (Brenner, S., & Lerner, R.A. (1992) *Proc Natl Acad Sci USA* 89, 5381-3; Kerr, J.M., Banville, S.C., & Zuckermann, R.N. (1993) *J Am Chem Soc* 115, 2529-2531; Needels, M.C., Jones, D.G., Tate, E.H., Heinkel, G.L., Kochersperger, L.M., Dower, W.J., Barrett, R.W., & Gallop, M.A. (1993) *Proc Natl Acad Sci USA* 90, 10700-4; Ohlmeyer, M.H., Swanson, R.N., Dillard, L.W., Reader, J.C., Asouline, G., Kobayashi, R., Wigler, M., & Still, W.C. (1993) *Proc Natl Acad Sci USA* 90, 10922-6). Additionally, biological libraries obtained from phage (Cwirla, S.E., Peters, E.A., Barrett, R.W., & Dower, W.J. (1990) *Proc Natl Acad Sci USA* 87, 6378-82; Smith, G.P. (1985) *Science* 228, 1315-1317) or bacterial (Cull, M.G., Miller, J.F., & Schatz, P.J. (1992) *Proc Natl*

*Acad Sci USA* 89, 1865-9; Goodson, R.J., Doyle, M.V., Kaufman, S.E., & Rosenberg, S. (1994) *Proc Natl Acad Sci USA* 91, 7129-33) systems, have influenced the development of synthetic combinatorial libraries.

The above methods all have certain general attributes that contribute to their success. There is virtually no limit to library size using these approaches, potentially increasing the likelihood of obtaining a "hit." Several of the methods listed above attempt to accommodate non-peptide libraries in inventive ways; however, many are burdened by the need for an increased synthetic effort and most deal with resin-bound substrates that may hinder substrate binding, particularly in cases where an active site contains a deep pocket. Furthermore, the selection processes themselves can be quite cumbersome due to the massive size of some libraries.

Therefore, there is needed a method for the characterization of members of a combinatorial library with particular and desirable properties.

## SUMMARY OF THE INVENTION

The present invention is drawn to a method for characterizing the members of a combinatorial library which bind to a domain of interest. The method utilizes affinity selection in combination with mass spectrometry. The method provides rapid and efficient screening and provides information on relative affinities, molecular weights, and structural identification of affinity-selected compounds. The method has broad applicability to combinatorial library screening and is useful for peptide as well as non-peptide libraries.

## BRIEF DESCRIPTION OF THE FIGURES

**Figure 1:** Displacement of the binding of Bolton-Hunter radiolabeled SH2 binding peptide by various displacers. Peptides were tested as described in methods for the ability to inhibit binding of a radiolabeled peptide to the C-terminal SH2 domain of PI 3-Kinase.

**Figure 2:** Comparison of various elution strategies for the separation of standard peptides on SH2-affinity columns. Standard peptides were run on the SH2-affinity column and eluted by various gradients as described in methods. Fractions of the gradient elution were run on reverse phase HPLC, quantitated using UV absorption at 210 nm, and plotted as the percent

of the total peptide eluted. Parallel runs were performed with a) KSCN, b) phenyl phosphate and c) pH elution.

**Figure 3:** ES-MS/MS spectrum of G($F_2$Pmp)(Nle)P(Nle)S-amide. The fragmentation pattern assignments are labeled according to the following: single letter designations refer to amino acids and internal fragments forming ammonium ions; b-ions refer to ions formed with the charge retained on the N-terminus upon cleavage at the peptide bond; a-ions relate to N-terminal fragment ions with cleavage occurring after the -carbon (28 Da less than corresponding b-ions); y-ions refer to ions formed with the charge retained on the C-terminus upon cleavage at the peptide bond (complimentary ions to corresponding b-ions). The sequence is designated above the b-ions in the mass spectrum.

**Figure 4:** ES-MS spectrum of the intact library defined as G($F_2$Pmp)$X_1$P$X_3$S-amide, where $F_2$Pmp is (phosphonodifluoromethyl)phenylalanine and X=19 of the 20 standard L-amino acids (minus cysteine). This library contains 361 components and the mass spectrum was obtained from 450 ng per peptide.

**Figure 5:** ES-MS spectrum of the library affinity selected components in a pH 3.5 affinity elution. Assignments are as designated as follows: sequences inferred for AM, SM, VM, DM, K/QM and YM; molecular weights validated for TM, NM and WM; and sequences confirmed for I/LM, EM, MM and FM (see Table I).

**Figure 6:** ES-MS/MS spectrum of the molecular ion at m/z 798.5 from the pH 3.5 elution. Assignments are as designated in Figure 3 and asterisked peaks indicate methionine-specific fragmentation ions. The sequence was determined to be that designated above the b-ions of the mass spectrum.

## DETAILED DESCRIPTION OF THE INVENTION

Methods for characterizing the members of a combinatorial library are provided. Generally, the members of a library are brought into contact with a domain of interest to allow for binding; i.e. the formation of a complex. After binding, the complex is separated from the unbound members of the library. The complexes are then treated to elute the bound library components. The eluted components are analyzed by mass spectrometry.

The combinatorial libraries which may be analyzed by the invention include peptide as well as non-peptide libraries including small molecule libraries. Peptide libraries may include

unnatural amino acids and amino acid mimics, such as norleucine, phosphotyrosine mimics, such as (phosphonodifluoromethyl) phenylalanine, and the like. See, for example, Kerr *et al.* (1993) *J Am Chem Soc 115*, 2529-31.

Methods for making combinatorial libraries are well known in the art. See, for example, Felder E.R. ( 1994) *Chimia 48*, 521-541; Gallop *et al.* (1994) *J Med Chem 37*, 1233-51; Houghten, R.A. (1993) *Trends Genet 9*, 235-9; Houghten *et al.*, (1991) *Nature 354*, 84-6; Lam *et al.*, (1991) *Nature 354*, 82-4; Carell *et al.*, (1995) *Chem Biol 3*, 171-183; Madden *et al.* (1994) *Perspectives in Drug Discovery and Design 2*, 269-85; Cwirla *et al.*, (1990) *Biochemistry 87*, 6378-82; Brenner, S. and Lerner, R.A. (1992) *Proc Natl Acad Sci USA 89*, 5381-3; Gordon *et al.* (1994) *J Medicinal Chemistry 37*, 1385-401; and, Lebl *et al.*, (1995) *Biopolymers 37*, 177-98.

The present method is useful for analyzing the binding to any domain or receptor. It finds use in the characterization of combinatorial libraries members which bind to high affinity acceptor molecules such as receptors, enzymes, antibodies, and the like. One can also screen for an activity of interest. Depending on the type of activity that is sought, a variety of selection schemes are possible based on the affinity purification of active members by absorption on a target receptor.

The method has a detection limit in the range of about 5 to about 50 mol. Thus, it is useful even when the compound of interest is present in limited amounts in a sample.

In carrying out the invention, the domain or receptor of interest may be bound or coupled to a support. Supports include resin beads, silica chips, agarose, and other solid supports. The domain or receptor can be bound by any method including but not limited to antibody binding, GST-glutathione binding, biotin-streptavidin, expression of the protein of interest as a recombinant protein fused to maltose binding protein, fusion of the component of interest with a marker peptide which recognizes and binds selectively to an affinity column, etc. Such methods are known in the art and kits for practicing the methods are commercially available. See, for example, Stammers *et al.* (1991) *FEBS Lett. 283*, 298-302; Herman *et al.* (1986) *Anal. Biochemistry 156*, 48; Smith *et al.* (1987) *FEBS Lett. 215*, 305; Kilmartin *et al.* (1982) *J. Cell. Biol. 93*, 576-82; Skinner *et al.* (1991) *J. Biol. Chem. 266*, 14163-6; Hopp *et al.* (1988)

*Bio/Technology 6*, 1204-10; H.M. Sassenfeld (1990) *TIBTECH 8*, 88-93; Hanke *et al. (1992) J. General Virology 73*, 653-60; Ellison and Hochstrasser (1991) *J. Biol. Chem. 267*, 21150-7; U.K. Pati (1992) *Gene 114*, 285-8; Wadzinski *et al.* (1992) *J. Biol. Chem. 267*, 16883-8; Field *et al.* (1988) *Mol. Cell. Biol. 8*, 2159-65; Gerard and Gerard (1990) *Biochemistry 29*, 9274-81; Ausselbergs *et al.* (1993) *Fibrinolysis 7*, 1-13; Hopp *et al.* (1988) *Biotechnology 6*, 1205-10; Blanar and Rutler (1992) *Science 256*, 1014-8; Lin and Morton (1991) *J. Org. Chem. 56*, 6850-6; Zastrow and Kobilka (1992) *J. Biol. Chem. 267*, 3530-8; Goldstein *et al.* (1992) *EMBO Jml. 11*, 0000-0000; Lim *et al.* (1990) *J. Infectious Disease, 162*, 1263-1269; Goldstein *et al.* (1992) *Virology 190*, 889-893; and the articles in *IBI FLAG Epitope* Vol. 1: No. 1, Sept. 1992; which disclosures are herein incorporated by reference.

Where these methods are used, the domain component comprises a site or region which is capable of binding to the support. For example, if GST-glutathione binding is used, the domain can be expressed as a GST-fusion protein which is coupled to glutathione-agarose. The method utilizes an affinity chromatography core comprising the support. That is, the use of specially designed absorbents or supports are used which depend upon specific biochemical reactions to selectively hold a macromolecule.

In this manner, the domain may be bound to the support prior to binding with the members of the combinatorial library. Alternatively, the domain may be unbound when mixed with the members of the combinatorial library. Where the domain is unbound, after the formation of a binding complex with the members of the combinatorial library, the complex can be brought into contact with a support to which the domain portion of the complex will bind. Generally, the domain will contain the attachment site prior to binding with members of the combinatorial library. However, it is recognized that the attachment site may be added after binding with the combinatorial library.

After binding to a support, the binding complexes can be separated from the unbound members of the combinatorial library by washing, for example through a column. "Binding complexes" is intended to signify the domain or receptor of interest bound to a member of the combinatorial library. The separation may vary depending on the binding method to the solid support. See the references listed above for binding methods. The elaboration of elution conditions and automated analysis are novel features of the present method.

After separating the bound complexes the present invention provides for characterization of the members by the physical parameters which drive the affinity of interaction and kinetics. In this manner, weakly bound members can be eluted and analyzed by mass spectrometry followed by the elution of the more strongly bound members. The members can thus be analyzed and identified by the weakly bound to the more strongly bound providing for a rapid method for rank-ordering of the members which bind to the domain of interest.

As noted the members can also be analyzed to determine the type of binding interactions between the library members and the domains. Several elution methods can be utilized to characterize the binding interactions of the members to the domain. Elution methods include but are not limited to phenylphosphate, phosphotyrosine mimic, or other suitable displacers, (for example displacing phosphopeptide which may be tagged for easy removal); chaotrope agents; pH elution; salt gradients; temperature gradients; organic solvents; selective denaturant; detergents; etc.

After elution, the members are analyzed by mass spectrometry. The mass spectra is useful to prove the identity of the members as well as to help establish the structure by providing molecular weight and formula of the members. In this manner, electrospray mass spectrometry (ES-MS) provides molecular weight information relating to combinatorial libraries and affinity-selected members. Electrospray tandem mass spectrometry (ES-MS/MS) provides the added advantage of sequence-specific determination of individual members that demonstrate high affinity binding. The library affinity selection-mass spectrometry (LAS-MS) of the invention is designed to exploit the attributes of solution phase libraries, affinity selection and mass spectrometry to study members of a combinatorial library that bind to domains of interest.

Methods for the analysis by mass spectrometry are generally known in the art. See, for example, Youngquist *et al.* (1995) *J Am Chem Soc 117*, 3900-06; Dunayevskiy *et al.* (1995) *Anal Chem 67*, 2906-15; Brummel *et al.* (1996) *Anal Chem 68*, 237-42; Metzger *et al.* (1994) *Analytical Biochemistry 219*, 261-277; Youngquist *et al.* (1994) *Rapid Comm Mass Spectrom 8*, 77-81; Till *et al.* (1994) *J Biol Chem 269*, 7423-8; and Brummel *et al.* (1994) *Science 264*, 399-402; which disclosures are herein incorporated by reference.

As indicated, the method finds use in characterizing combinatorial libraries. Any library including peptides, peptides containing non-natural amino acids, non-peptides, small molecules can be analyzed for binding to any domain of interest. Domains include proteins, enzymes, receptors, binding domains such as SH2 domains, and the like. Thus, the method finds use in understanding protein-protein interactions, enzymes-receptor interactions, etc., and in identifying individual compounds having specific biological activities of interest, antibody recognition sequences, antagonists, antigenic determinants, bioactive peptides, drug discovery and the like.

In a preferred embodiment of the invention, the LAS-MS method of the invention is designed to exploit the attributes of solution phase libraries, affinity selection and mass spectrometry to study peptides containing non-natural amino acids that bind to the Src homology 2 (SH2) domain, particularly to the SH2 domain of phosphatidylinositol 3-kinase (PI 3-Kinase). The LAS-MS technique is designed to rapidly screen for drugs with the potential to block signal transduction processes and has the additional advantage of allowing rapid rank-ordering of substrates that bind to SH2 in this system.

SH2 domains are an important protein motif (Pawson, T. (1995) *Nature* 373, 573-80; Schlessinger, J. (1994) *Curr Opin Genet Dev* 4, 25-30; Stahl, M.L., Ferenz, C.R., Kelleher, K.L., Kriz, R.W., & Knopf, J.L. (1988) *Nature* 332, 269-72), found particularly in cell systems relate to cell growth, proliferation and differentiation. SH2 domains functionally bind phosphotyrosine-containing sequences and are involved in protein-protein interactions. Of 50 or more SH2 comains identified (Klippel, A., Escobedo, J.A., Fantl, W.J., & Williams, L.T. (1992) *Mol Cell Biol* 12, 1451-9; Moran, M.F., Koch, C.A., Anderson, D., Ellis, C., England, L., Martin, G.S., & Pawson, T. (1990) *Proc Natl Acad Sci USA* 87, 8622-6; Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., & *et al.* (1993) *Cell* 72, 767-78; Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., & *et al.* (1994) *Mol Cell Biol* 14, 2777-85), short peptides have been shown to suffice for binding, with the amino acids immediately surrounding the phosphotyrosine serving to define specificity for each unique system. In the case of PI 3-Kinase (Herbst, J.J., Andrews, G., Contillo, L., Lamphere, L., Gardner, J., Lienhard, G.E., & Gibbs, E.M. (1994)

*Biochemistry* 33, 9376-81; Klippel, A., Escobedo, J.A., Fantl, W.J., & Williams, L.T. (1992) *Mol Cell Biol* 12, 1451-9; Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M.J., Hiles, I., Pawson, T., & Waterfield, M.D. (1993) *Mol Cell Biol* 13, 3567-76), the SH2 domain requires a methionine as the third residue on the C-terminal side of the phosphotyrosine (referred to as the +3 or $X_3$ residue), as originally established using peptides derived from the natural growth factor receptors (Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., & *et al.* (1993) *Cell* 72, 767-78; Songyang, Z., Shoelson, S.E., McGlade, J., Olivier, P., Pawson, T., Bustelo, X.R., Barbacid, M., Sabe, H., Hanafusa, H., Yi, T., & *et al.* (1994) *Mol Cell Biol* 14, 2777-85).

Several phosphotyrosine mimics have been identified that are able to maintain high binding affinity to SH2 domains (Burke, T.R., Jr., Smyth, M.S., Otaka, A., Nomizu, M., Roller, P.P., Wolf, G., Case, R., & Shoelson, S.E. (1994) *Biochemistry* 33, 6490-4; Ye, B., Akamatsu, M., Shoelson, S.E., Wolf, G., Giorgetti-Peraldi, S., Yan, X., Roller, P.P., & Burke, T.R., Jr. (1995) *J Med Chem* 38, 4270-5). These mimics offer the advantage of a stabilized, phosphatase-resistant substitution and, additionally, may bring about more favorable cellular permeability properties as compared to phosphotyrosine. For example, methylphosphonate analogs approach the affinity of phosphotyrosine for SH2-domains (Burke, T.R., Jr., Smyth, M.S., Otaka, A., Nomizu, M., Roller, P.P., Wolf, G., Case, R., & Shoelson, S.E. (1994) *Biochemistry* 33, 6490-4).

Coupling the affinity selection process with electrospray mass spectrometry (ES-MS) and tandem mass spectrometry (ES-MS/MS), in the present method, is an effective means to identify and characterize peptides within a combinatorial library that bind to the SH2 domain of PI 3-Kinase. The practicality of this method is demonstrated whereby a simplified approach can be utilized without the need for a sophisticated chromatographic set-up or a dedicated/modified mass spectrometric system. Recently, mass spectrometry has been used to characterize the integrity of combinatorial libraries (Brummel, C.L., Lee, I.N., Zhou, Y., Benkovic, S.J., & Winograd, N. (1994) *Science* 264, 399-402; Dunayevskiy, Y., Vouros, P., Carell, T., Wintner, E.A., & Rebek, J.J. (1995) *Anal Chem* 67, 2906-2915; Kaur, S., Huebner, V., Tang, D., McGuire, L., Drummond, R., Csetjey, J., Stratton-Thomas, J., Rosenberg, S., Figliozzi, G., Banville, S., Zuckermann, R., & Dollinger, G. (1995) in *43rd*

*ASMS Conference on Mass Spectrometry and Allied Topics* pp 30, Atlanta, GA.; Metzger, J.W., Kempter, C., Wiesmuller, K.H., & Jung, G. (1994) *Anal Biochem* 219, 261-77). Where ES-MS provides molecular weight information, ES-MS/MS provides molecular (sequence) identification to resolve the redundancies in molecular weights associated with many libraries. In addition, mass spectrometry is readily applicable to peptido-mimics and non-peptidic, small molecule libraries.

The following examples are offered by way of illustration and not by way of limitation.

## EXPERIMENTAL

### *Chemicals*

Glutathione-Agarose was purchased from Pharmacia (Piscataway, NJ). Other chemicals were purchased from Sigma (St. Louis, MO). The standard peptides were synthesized by the FMOC solid phase method. The combinatorial library was purchased from Multiple Peptide Systems (San Diego, CA) and was defined as $G(F_2Pmp)X_1PX_3S$-amide, where $F_2Pmp$ is (phosphonodifluoromethyl)phenylalanine and X=19 of the 20 standard L-amino acids (minus cysteine).

### *GST-SH2 Purification*

A GST-SH2 fusion protein derived from the C-terminal SH2 domain (residues 617-722) of human PI 3-Kinase (Panayotou, G., Gish, G., End, P., Truong, O., Gout, I., Dhand, R., Fry, M.J., Hiles, I., Pawson, T., & Waterfield, M.D. (1993) *Mol Cell Biol* 13, 3567-76) was subcloned and grown in an *E. coli* expression system. Colonies were grown overnight and used to seed one liter preparations. The larger preparations were then grown to an O.D. of 0.6-0.8 and induced for 3 hours with 1 mM IPTG. In a typical preparation, 2 grams of cell pellet was suspended in 12 ml of 1% Triton in phosphate buffered saline (PBS), sonicated 3x20 seconds and centrifuged at 8000 rpm for 20 min using a SS-34 rotor. The supernatant was mixed with 10 mL of Glutathione-Agarose to enable the GST-SH2 to bind to the matrix. The mixture was gently shaken at 4 C for 30 min, loaded onto an empty column and washed

with 200 ml of PBS. GST-SH2 was eluted from the column using 50 mM glutathione in 100 mM $NH_4HCO_3$, pH 8.6.

*SH2 Affinity Chromatograph*

The GST-SH2 fusion protein (12 mg) was coupled to 1 mL Glutathione-Agarose. The gel was pre-equilibrated with 50 mM ammonium acetate buffer pH 7.5 (equilibrating buffer). For the batch method of library analysis, the library (2.5 mg) was dissolved in 1 mL of equilibrating buffer and mixed with 1 mL of beads at 4 C for 1 hr. The SH2-GST Glutathione-Agarose matrix was washed 3 times with 10 bed volumes of equilibration buffer and the weakly bound peptides were removed by 3 successive washes with 1 bed volume of 1 M ammonium acetate buffer pH 7.5. More strongly bound library components were eluted successively with 100 mM triethylamine acetate (TA) buffer, pH 4.5; 100 mM TA buffer, pH 3.5; and 0.5% trifluoroacetic acid (TFA). Each step entailed three consecutive washes with two bed volumes of elution buffer and the eluted peptides were monitored by either ES-MS or ES-MS/MS (see below).

In other applications, a column-based gradient elution was used employing one of three linear gradients over 80 min at a flow rate of 0.5 mL/min. The pH gradient was from pH 7.5 (50 mM ammonium acetate) to 3.5 (100 mM triethylamine acetate), the chaotrope (KSCN) gradient ranged from 0 to 1 M and the displacer (sodium phenyl phosphate) gradient from 0 to 0.25 M. The eluted peptides were monitored by reverse phase high performance liquid chromatography (HPLC).

*Other Methods*

An SH2 binding assay was performed using a Bolton-Hunter labeled phosphopeptide essentially as described by Piccione *et al.* (Piccione, E., Case, R.D., Domchek, S.M., Hu, P., Chaudhuri, M., Backer, J.M., Schlessinger, J., & Shoelson, S.E. (1993) *Biochemistry* 32, 3197-202). Peptides were analyzed using a Biorad (Richmond, CA) 250 x 4.6 mm RP 318 reverse phase HPLC column with a 30 min linear gradient from 5% B buffer (90% acetonitrile (AcN), 0.09% TFA)/95% A buffer (5% AcN, 0.09% TFA) to 25% B/75% A. HPLC was

performed at room temperature at 1 mL/min and was monitored for absorption at 210 nm with a Waters 490 UV detector.

*Mass Spectrometry*

Electrospray mass spectrometry and tandem mass spectrometry was performed on a PE Sciex API III triple quadruple mass spectrometer (Concord, Ontario). A Shimatzu (Kyoto, Japan) LC-10 HPLC was used to deliver a flow rate of 30 L/min directly to the electrospray interface. Samples were concentrated using a Michrom (Auburn, CA) small molecule trap placed in-line with the HPLC system and washed with 2:98 A:B (A: AcN and B: 0.1% TFA) to remove interfering salts. Trapping of peptides was maintained using a 2:98 A:B ratio in the LC system, and the elution of peptides was facilitated by increasing the A:B ratio to 80:20 for direct infusion of peptides into the mass spectrometer.

All ES-MS and ES-MS/MS analyses were carried out in the positive ion mode. Typical ES-MS conditions utilized an ionization voltage of 5 kV and orifice voltage of 70 V. Purified air was used as a nebulization gas (40 psi) and ultrapure nitrogen was used as a curtain gas at a flow rate of 1 L/min. ES-MS/MS was carried out using ultrapure argon at 260 x $10^{12}$ molecules/$cm^2$ collision gas thickness and a 29 eV collision energy.

The intact peptide library was dissolved in a 2:98 ratio (A:B) and analyzed from a 0.5 mg portion of the original sample. All ES-MS spectra were obtained using multi-channel analyzer (MCA) mode and scanning the mass range 630 - 950 m/z in 0.1 m/z increments and with a 1 msec dwell time. The calculated amount of each peptide utilized to obtain the ES-MS and ES-MS/MS spectra from the intact library was 450 ng.

ES-MS/MS experiments were carried out using a scan range of 50 - (parent ion + *ca.* 10) m/z units in 0.1 m/z increments and a 1 msec dwell time. Limitations to MS/MS on a triple quadruple instrument are due to practical sequence interpretation limits for compounds *ca.* > 2000 Da and to the instrument's inability to distinguish between amino acids isoleucine and leucine (I/L; MW = 113 Da) or lysine and glutamine (K/Q; MW = 128 Da). The amino acid ambiguity can be overcome by eliminating one of the redundant residues from the library or by substituting an isotopic along for one of the two.

*Results*

*Pilot Experiments*

Purified human PI 3-Kinase SH2 domain was produced as a fusion protein with glutathione-S-transferase (GST). The SH2 fusion protein was immobilized to Glutathione-Agarose and used to screen a biased combinatorial library. Pilot experiments with model compounds were performed to test this affinity isolation technique. The four standard compounds utilized were: GY(Nle)P(Nle)S-amide; G(pY)(Nle)P(Nle)S-amide; G($F_2$Pmp)(Nle)P(Nle)S-amide; and GNGDS(pY)MPMPKS-amide (Nle = norleucine; Py = phosphotyrosine; and $F_2$Pmp = (phosphonodifluoromethyl(phenylalanine). Three gradient elution methods (phenylphosphate displacer, chaotrope agent and pH elution) were evaluated for separating these compounds on the SH2 affinity column. Binding assays of these four standard compounds revealed the $IC_{50}$ values shown in Figure 1.

All three gradient procedures selectively released the standards from the SH2 affinity system (Figure 2). However, the order of elution differed among the various methods, most likely reflecting the differing mechanisms by which the elution agents were disrupting the binding interactions. For example, GNGDS(pY)MPMPKS-amide eluted late in the presence of the displacer gradient, but relatively early when a chaotrope gradient was used. This compound has the highest affinity for the SH2 domain (Figure 1), but eluted earlier than the other standards when analyzed by reverse phase HPLC (not shown), due to its highly charged sequence. The pH gradient elution was suitable for subsequent coupling of this assay with mass spectrometry.

A software program was developed (see Appendix A) in order to predict the molecular weight distribution of the components of the tested library. The high level logic of the program is included as Appendix B. For this example, redundancies of two or more components exist because the same group of building blocks were used for positions $X_1$ and $X_3$. For example, each pair of variable residues such as SD and DS will have the same molecular weight. Although ES-MS analysis of the library will not distinguish between these redundancies, ES-MS/MS analysis can, due to its sequencing ability.

The ES-MS system was tested with a standard compound, G($F_2$Pmp)(Nle)P(Nle)S-amide, to ensure the effectiveness of the peptide trap for the modified peptides present in the library. This compound was retained by the trap and resulted in the expected molecular ion {$(M+H+)^+$} of m/z 762.5 ($MW_{calc}$ = 761.4). ES-MS/MS analysis was performed to further explore the capabilities of the system for the present application, and to probe for molecular identity and sequence information.

The ES-MS/MS spectrum of the ion at m/z 762.5 displayed an informative fragmentation pattern (Figure 3). The "b" ions (N-terminal ions formed upon cleavage of the peptide bond) formed a clear pattern starting with the G($F_2$Pmp) ion at m/z 334.4. From this ion the remaining amino acids could be detected in the correct order: (Nle)P(Nle)S. Each b-ion had a paired "a" ion at -28 Da (-CO). In addition, proline residues are known to be very liable in MS/MS, where the cleavage takes place at the N-terminal side (so-called "internal ions"). This was the case for the above peptide, where internal fragment ions for P(Nle) and P(Nle)S were detected. These fragmentation patterns were valuable for the characterization of peptides in the library so that the order of the $X_1$ and $X_3$ amino acids could be determined via the b-ions {$X_1$ after G($F_2$Pmp)} and the internal ions {$X_3$ after P}.

*Characterization of a Combinatorial Library by MS*

The combinatorial library was defined as: G($F_2$Pmp)$X_1$P$X_3$S-amide, where X = 19 of the 20 standard L-amino acids (minus cysteine) and, therefore, consisted of 361 ($19^2$) compounds. $F_2$Pmp ((phosphonodifluoromethyl(phenylalanine) was utilized as a high affinity phosphotyrosine mimic, as discussed by Burke *et al.* (Burke, T.R., Jr., Smyth, M.S., Otaka, A., Nomizu, M., Roller, P.P., Wolf, G., Case, R., & Shoelson, S.E. (1994) *Biochemistry* 33, 6490-4). ES-MS analysis of the intact library revealed the diversity of the 361 components of the library, as shown in Figure 4. The distribution of the peptides properly reflected the expected statistical distribution of molecular weight combinations of amino acids. This mass spectrum also resembled an ES-MS spectrum reported for a 576 component peptoid mixture (Kaur, S., Huebner, V., Tang, D., McGuire, L., Drummond, R., Csetjey, J., Stratton-Thomas, J., Rosenberg, S., Figliozzi, G., Banville, S., Zuckermann, R., & Dollinger, G. (1995) in *43rd ASMS Conference on Mass Spectrometry and Allied Topics* pp 30, Atlanta, GA.).

An ES-MS/MS spectrum was obtained of the ion at m/z 796.6 from the total library mixture. Knowing the four fixed residues in the library, this molecular ion should represent eight distinct peptides containing $X_1$ and $X_3$ combinations: EM, PY, IF/LF, ME, YP, FI/FL. Inspection of the MS/MS spectrum (not shown), although complicated by the mixture, indicated the presence of all of the above combinations.

*Library Affinity Selection*

Affinity selection of the library was performed using a batch procedure. Five pH elution steps were evaluated: unbound, 1M salt/pH 7.5 wash; pH 4.5 wash; pH 3.5 and TFA elutions, respectively. The unbound resin supernatant, taken prior to any washing of the SH2-peptide mixture, contained few peptides; however, many of those detected matched MW's of glycine-containing combinations. The 1M salt/pH 7.5 elution contained the majority of the library components, that is those compounds weakly bound by SH2. The pH 4.5 elution contained fewer peptides, presumably those moderately bound.

As seen in Figure 5, a small subset of components were detected in the pH 3.5 elution. Virtually all of these compounds could be assigned as methionine-containing peptides. Table I lists the molecular ions found over an intensity threshold value of 15 percent in the pH 3.5 elution, as well as the combinations associated with each molecular weight. Of four possible unambiguous methionine-containing combinations (TM, NM, HM and WM), three such ions (TM, NM and WM) were detected in the pH 3.5 elution. Additionally, four predicted methionine-containing peptides were not detected in the pH 3.5 elution (GM, PM, HM and RM). Of these, the GM peptide was found in the pH 7.5 and 4.5 elutions, the PM and HM peptides were found in the pH 4.5 elution and the RM peptide was seen but was below the detection threshold level (15%) in the pH 3.5 elution. Three peptides were detected (m/z 735.7, 779.7 and 821.0) that did not match a methionine-containing molecular weight. Although the recurrence of methionine in the strongly bound peptides is readily discernible from the data in Table I, the identity and order of residues in positions $X_1$ and $X_3$ must still be probed by ES-MS/MS.

The most abundant peptide in the pH 3.5 elution was m/z 798.5, which showed a significant increase from previous elutions and was also found in the TFA elution. This MW matches that for MM, FD and VY combinations. ES-MS/MS analysis of this ion revealed that only the MM combination was present in the pH 3.5 elution (Figure 6). Given the fact that this ion is also found in the TFA elution, the MM combination can be qualitatively be assigned as a strong binder to SH2.

To probe other redundancies listed in Table I, ES-MS/MS spectra were obtained for the same ion, m/z 780.5, over the pH elutions of 7.5, 4.5 and 3.5. Possible sequences of $X_1$ and $X_3$ for this ion are: IM/LM, ED, PF, MI/ML, De and FP. Changes in the fragmentation patterns can be detected with decreasing pH. Most importantly, the position of the methionine can be seen to favor position $X_3$ at pH 3.5 for the (I/L)M combination. In addition, the peptide at m/z 780.5 was detected in the TFA elution. This implies that the (I/L)M combination and the MM combinations are the strongest binders to SH2, consistent with previously reported data (Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., & *et al.* (1993) *Cell* 72, 767-78). Other ES-MS/MS analyses obtained {e.g. for ions at m/z 796.3 and 814.3 (see Table I)} detected the EM and FM sequences (in order of positions $X_1$ and $X_3$), revealing that methionine predominates at position $X_3$ in the pH 3.5 elution. Further software development has been performed in order to facilitate data analysis and data in profiles. This program is included as Appendix C. The high level logic is provided as Appendix D.

*Discussion*

The present results have demonstrated that screening combinatorial libraries in conjunction with ES-MS and ES-MS/MS detection is a feasible approach.

SH2 domains have been extensively documented by various biophysical approaches, and the forces responsible for high affinity phosphopeptide interaction are well established. The phosphotyrosine group represents the greatest energy of interaction, while residues (chiefly) immediately C-terminal contribute to binding energy and specificity. In the case of PI 3-Kinase, the pioneering work by Cantley (Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., & *et al.*

(1993) *Cell* 72, 767-78) demonstrated the importance of a hydrophobic residue in the +1 site and a methionine in the +3 site. This concept was confirmed using the LAS-MS approach and extend these observations to show: 1) that different combinations of such groups have widely different affinities and can be separated by judicious use of elution conditions; and 2) that the same principle applies even with libraries based upon unnatural phosphotyrosine mimics.

Various methods of elution have been exploited for affinity chromatography of proteins and each method has unique advantages and disadvantages. Three possible elution techniques were compared here using model peptides. Each technique showed an individual character reflecting the physical parameter utilized by each method. Displacer elution using phenylphosphate, as originally used by Cantely (Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J., & *et al.* (1993) *Cell* 72, 767-78), is a mild method which released components in a manner reflective of the relative affinities at the phosphate binding pocket. Acid elution was likewise effective and its elution profile relates to the titration of acidic amino acids in the SH2 domain as well as the titration of phosphonate and phosphate groups. Chaotrope elution is based upon displacement of hydrophobic interactions, a parameter that can vary enormously within a given library. While there is no reason to suggest that one method is preferable to another for the screening of libraries, it is clear that different rank-orders of potency are possible and may carry useful information reflective of the active site and the nature of the binding interactions.

LAS-MS detection of library components has proven to be a powerful technique for sequence-specific determination of high affinity compounds. ES-MS is a rapid, sensitive technique that has great potential for this type of application due to its molecular specificity, while tandem instruments offer specific structural elucidation. Although there are ambiguities determining certain residues such as I/L and K/Q with a triple quadruple mass spectrometer, these can be overcome to a degree by eliminating one of the residues from a given library. Other obstacles may present themselves due to the redundant molecular weights that are often present in the structural combinations contained in a peptide library. These too can be overcome by planning a theoretical library that will not contain such redundancies, potentially designed on a experimental manipulation of data could prove to be extremely useful in both

the planning of libraries and the potential automation of data analysis. Another approach would be to re-synthesize individual components of redundant molecular weight for further testing. Furthermore, as combinatorial libraries increasingly move toward the use of unique building blocks (Gordon, E.M., Barrett, R.W., Dower, W.J., Fodor, S.P., & Gallop, M.A. (1994) *J Med Chem* 37, 1385-401), these redundancies will become non-issues.

The analysis of intact combinatorial libraries should be a priority before carrying out bioassays, a practical consideration that is often overlooked. ES-MS allows for data to be obtained from which the quality of the library can be assessed. If necessary, a more extensive analysis can be performed on the library using ES-MS/MS. Recent reports addressing such issues (Dunayevskiy, Y., Vouros, P., Carell, T., Wintner, E.A., & Rebek, J.J. (1995) *Anal Chem* 67, 2906-2915; Metzger, J.W., Kempter, C., Wiesmuller, K.H., & Jung, G. (1994) *Anal Biochem* 219, 261-77) are extremely informative and essential to the field of combinatorial chemistry.

The strength of ES-MS/MS detection used in conjunction with library affinity selection lies in its ability to determine the identity of molecular ions of interest. The information obtained directly correlates to individual molecules rather than to a selected mixture. For example, Edman sequencing of affinity-selected peptides is only able to sequence pooled, high affinity components as a mixture. When one amino acid is detected at a given position more frequently than others, it is shown to favor that position. ES-MS/MS can extend relative positional frequency determinations to include the precise pairing and sequence order of amino acids for individual components detected from library affinity selection. Even when molecular weight redundancies exist for paired combinations (for example, see Table I), ES-MS/MS can determine the amino acid sequence and eliminate the other redundancies concurrently.

The present work demonstrates how library affinity selection with mass spectrometric detection is possible utilizing the basic resources available in a pharmaceutical company, while addressing the many practical considerations that have been identified in the field of combinatorial chemistry. This methodology is ideal for the analysis of soluble, unbound library components, which can be of unlimited design with respect to small-molecule libraries. Significant aspects of the current method include: the simplicity of the experimental design

using a peptide trap; the ability to analyze the integrity of an intact combinatorial library; the practicality of a well-planned library related to the affinity system; the confirmation that $F_2$Pmp can replace phosphotyrosine for SH2 binding; the selectivity of the method with the possibility for ranking; and the key advantage of ES-MS/MS for amino acid sequencing, applied to non-standard amino acids.

All publications and patent applications mentioned in the specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

**Table I: Library Affinity Selected Compounds Detected by ES-MS (pH 3.5 Elution)** Tabulations are based upon molecular ions {$(M+H)^+$} and their intensities detected by ES-MS above a threshold of 15 percent in the pH 3.5 affinity elution (see Figure 5). The asterisked ions under methionine-containing combinations are assigned as isotope peaks from the corresponding monoisotopic ions. All combinations shown contain redundancies related to sequence order. MW validated ions were derived from pairs with non-redundant molecular weights. Sequence validated ions were confirmed using ES-MS/MS to determine amino acid position and order.

**TABLE I**

| Found at pH 3.5 (M+H)+ | Intensity (x1000) | Met Combinations | Other Combinations | MW validated (ES-MS) | Sequence validated (ES-MS/MS) |
|---|---|---|---|---|---|
| 735.7 | 340 | | AK/Q | | |
| 738.4 | 470 | AM | TT | | |
| 754.5 | 640 | SM | AF | | |
| 766.5 | 360 | VM | TE,DD | | |
| 767.6 | 260 | * | | | |
| 768.5 | 530 | TM | | YES | |
| 779.7 | 510 | | NE,K/QD, GW,SR | | |
| 780.5 | 1270 | I/LM | ED,PF | | YES |
| 781.5 | 630 | NM | | YES | |
| 782.4 | 580 | DM | VF | | |
| 783.5 | 280 | * | | | |
| 795.5 | 490 | K/QM | | | |
| 796.5 | 830 | EM | PY,I/LF | | YES |
| 797.6 | 630 | * | | | |
| 798.5 | 1700 | MM | FD,VY | | YES |
| 799.5 | 640 | * | | | |
| 800.4 | 370 | * | | | |
| 814.5 | 840 | FM | DY | | YES |
| 815.5 | 450 | * | | | |
| 821.0 | 290 | | ER,VW | | |
| 830.5 | 340 | YM | FF | | |
| 853.5 | 440 | WM | | YES | |

What is claimed is:

1. A method for characterizing the members of a combinatorial library which bind to a domain of interest, said method comprising:

- mixing said domain of interest with a combinatorial library to allow binding of members with said domain to form binding complexes;
- separating complexes from unbound members;
- eluting the bound members of said complexes; and,
- analyzing said members by mass spectrometry.

2. The method of claim 1, wherein said domain is bound to a matrix.

3. The method of claim 2, wherein said binding is by GST-Glutathione binding, antibody antigen binding, or biotin- streptavidin binding.

4. The method of claim 3, wherein said binding is by GST- Glutathione binding.

5. The method of claim 2, wherein said matrix is selected from agarose, resin beads, and silica chips.

6. The method of claim 5 wherein said matrix is agarose.

7. The method of claim 1, wherein said combinatorial library is a peptide library.

8. The method of claim 7, wherein said peptide library contains unnatural amino acid mimics.

9. The method of claim 8, wherein said unnatural amino acid mimics are selected from norleucine and phosphotyrosine mimics.

10. The method of claim 1, wherein said combinatorial library is a non-peptide library.

11. The method of claim 1, wherein said eluting step is accomplished by phenylphosphate displacers, chaotrope agents, or pH elution.

12. The method of claim 11, wherein said eluting step is accomplished by phenylphosphate displacers.

13. The method of claim 11, wherein said eluting step is accomplished by chaotrope agents.

14. The method of claim 11, wherein said eluting step is accomplished by pH elution.

15. The method of claim 11, wherein said eluting step provides for determination of binding affinity and modes of binding interactions.

16. The method of claim 15, wherein said determination provides for successive elutions to remove weakly bound members followed by more strongly bound members.

17. The method of claim 1, wherein said mass spectrometry comprises electrospray mass spectrometry.

18. The method of claim 17, wherein said mass spectrometry further comprises electrospray tandem mass spectrometry.

19. A method for characterizing the members of a combinatorial library which bind to a Src homology 2 (SH2) domain, said method comprising:

- mixing said SH2 domain with a combinatorial library to allow binding of members with said domain to form binding complexes;
- separating complexes from unbound members;
- eluting the bound members of said complexes; and,
- analyzing said members by mass spectrometry.

20. The method of claim 19, wherein said SH2 domain is a domain of phosphatidylinositol 3-kinase (PI 3-Kinase).

21. The method of claim 19, wherein said domain is bound to a matrix.

22. The method of claim 21, wherein said binding is by GST-Glutathione binding, antibody binding, or biotin-streptavidin binding.

23. The method of claim 22, wherein said binding is by GST- Glutathione binding.

24. The method of claim 21, wherein said matrix is agarose.

25. The method of claim 19, wherein said combinatorial library is a peptide library.

26. The method of claim 19, wherein said combinatorial library is a non-peptide library.

27. The method of claim 19, wherein said eluting step is accomplished by phenylphosphate displacers, chaotrope agents, or pH elution.

28. The method of claim 27, wherein said eluting step is accomplished by phenylphosphate displacers.

29. The method of claim 27, wherein said eluting step is accomplished by chaotrope agents.

30. The method of claim 27, wherein said eluting step is accomplished by pH elution.

31. The method of claim 27, wherein said eluting step provides for determination of the relative strength of binding affinity.

32. The method of claim 31, wherein said determination provides for successive elutions to remove weakly bound members followed by more strongly bound members.

33. The method of claim 19, wherein said mass spectrometry comprises electrospray mass spectrometry.

34. The method of claim 33, wherein said mass spectrometry further comprises electrospray tandem mass spectrometry.

35. A computer readable medium having a program for the identification of compounds in a combinatorial library, the program comprising:

- means for creating a list of all possible members of said library;
- means for eliminating members that differ from another member only in the order of substituted parts; and
- means for creating a table with one row for each molecular weight occurring among the library members.

36. The medium of claim 35 disposed within a portable diskette.

37. The medium of claim 35 disposed within a computer and interconnected with a processor within said computer.

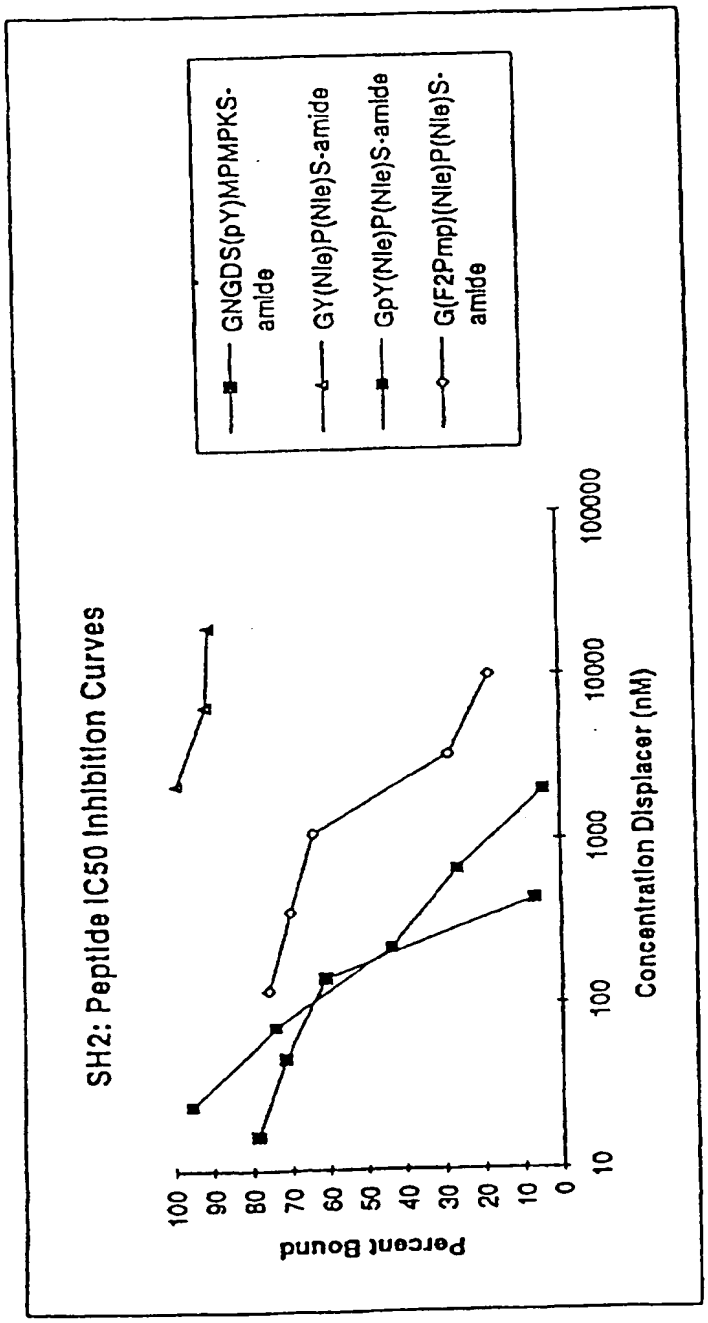
38. A computer readable medium having a program for the identification of compounds in a combinatorial library, the program comprising:

- means for creating a list of a known library of mass peaks; and
- means for identifying unknown mass peaks by mapping said mass peaks against said list of known mass peaks.

39. The medium of claim 38 disposed within a portable diskette.

40. The medium of claim 38 disposed within a computer and interconnected with a processor within said computer.

Figure 1




| | IC50 (nM) Ave. | SD |
|---|---|---|
| GNGDS(pY)MPMPKS-amide | 186 | 53 |
| GY(Nle)P(Nle)S-amide | >20000 | |
| GpY(Nle)P(Nle)S-amide | 202 | 24 |
| G(F2Pmp)(Nle)P(Nle)S-amide | 2881 | 2958 |

Figure 2a

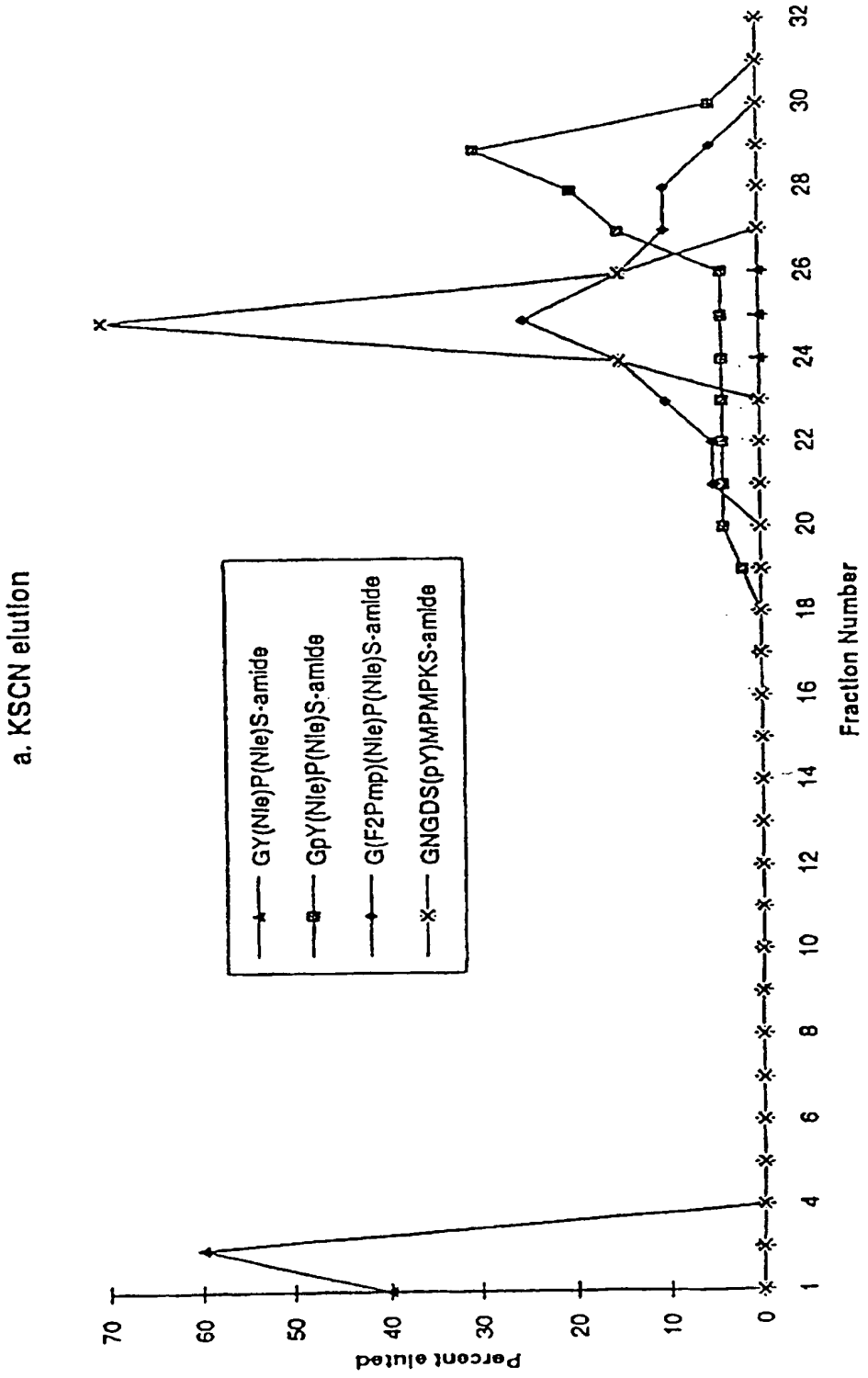


Figure 2b

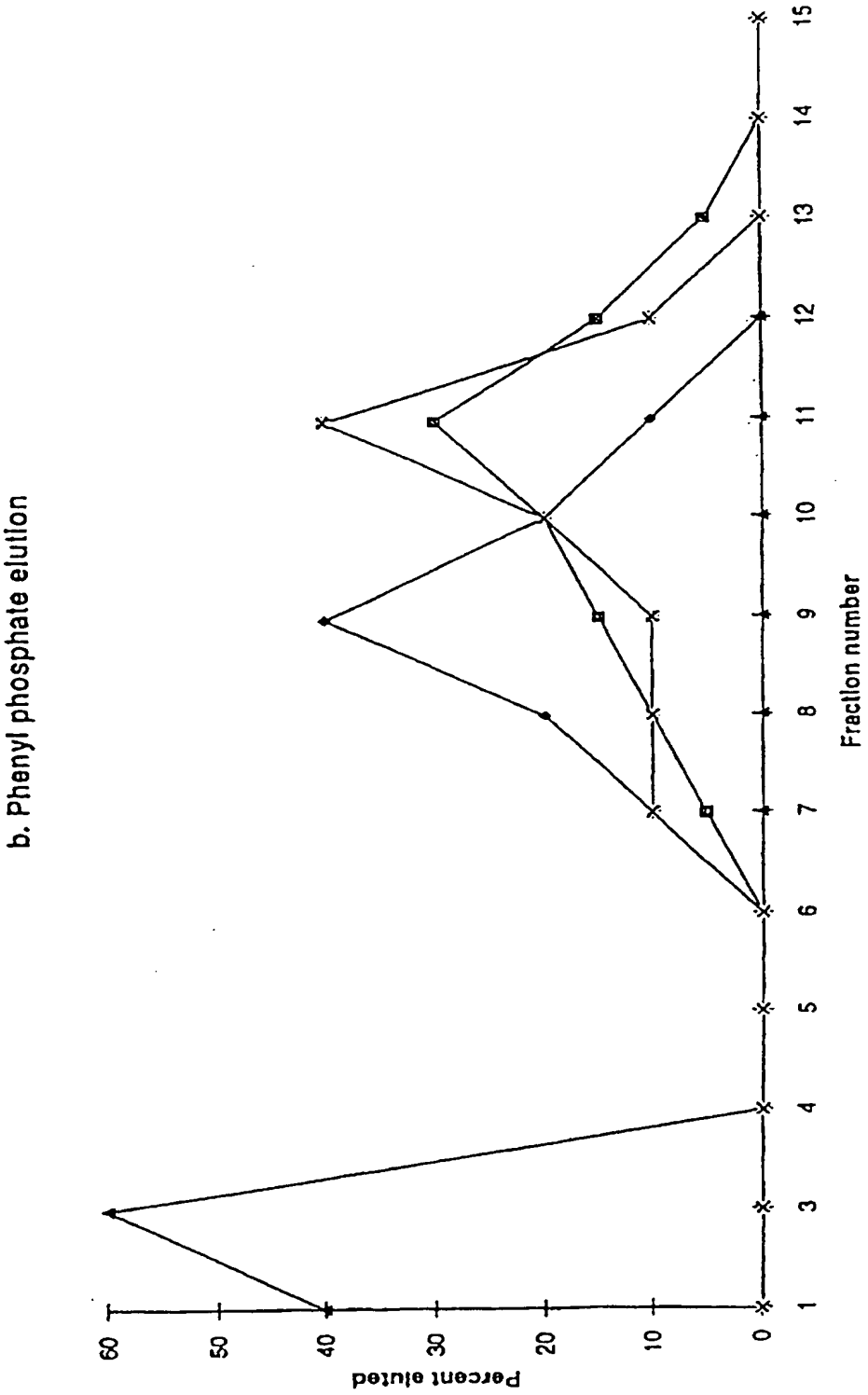

b. Phenyl phosphate elution
Percent eluted
0
10
20
30
40
50
60
Fraction number
1
3
4
5
6
7
8
9
10
11
12
13
14
15

Figure 2c

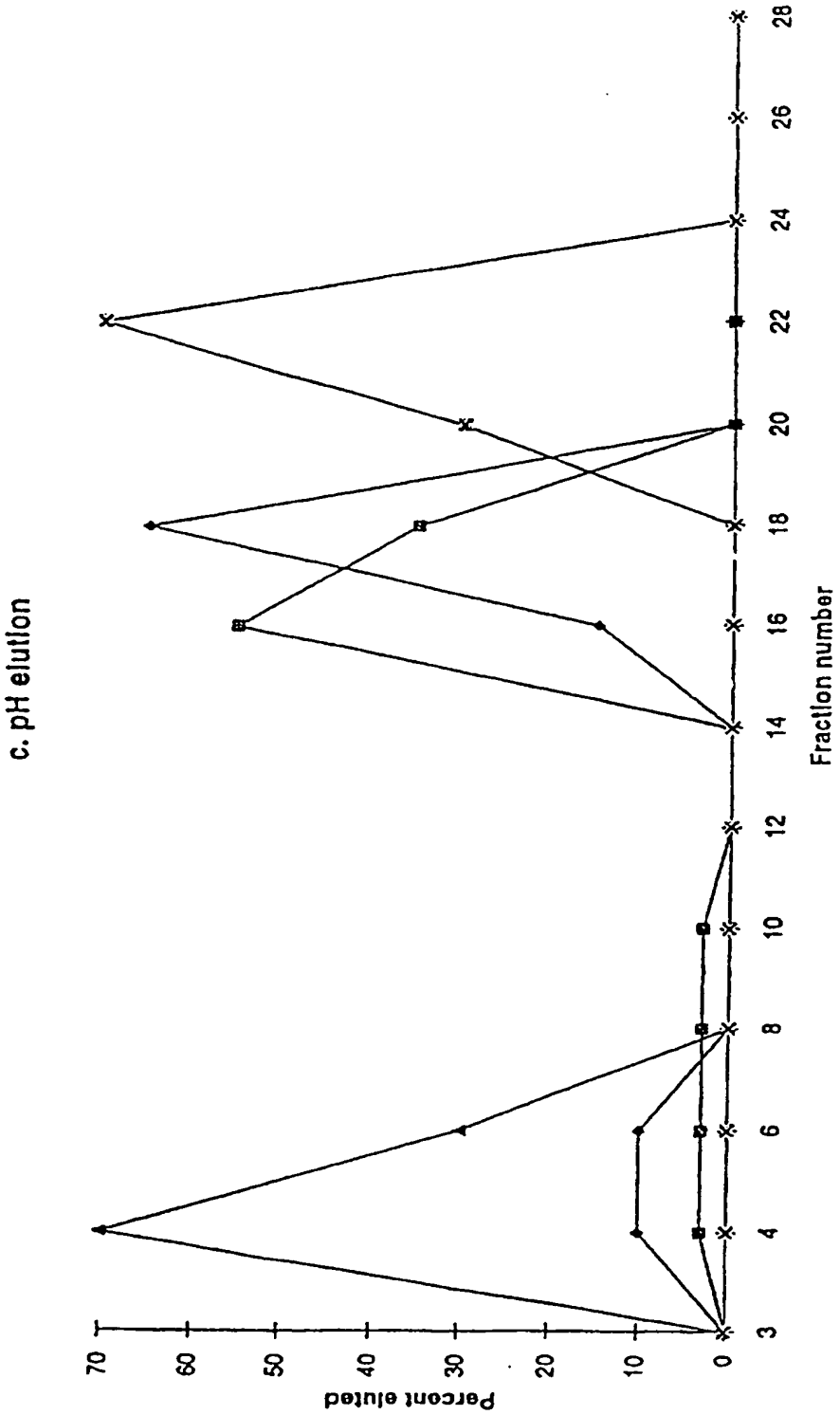


Figure 3

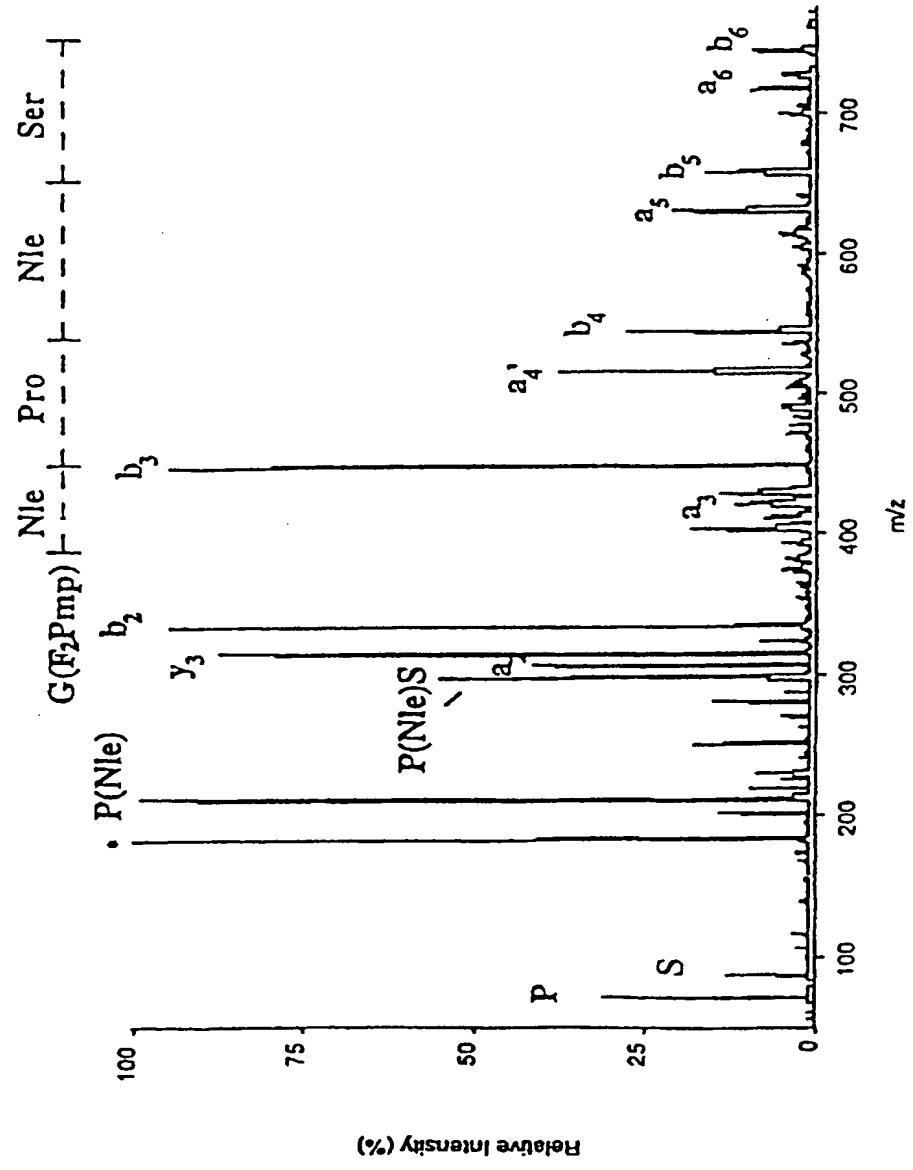

Ser
Nle
Pro
Nle
G(F2Pmp)
P(Nle)
b2
y3
P(Nle)S
a1
b3
a3
a4'
b4
a5
b5
a6
b6
P
S
100
75
50
25
0
100
200
300
400
500
600
700
m/z
Relative Intensity (%)

Figure 4

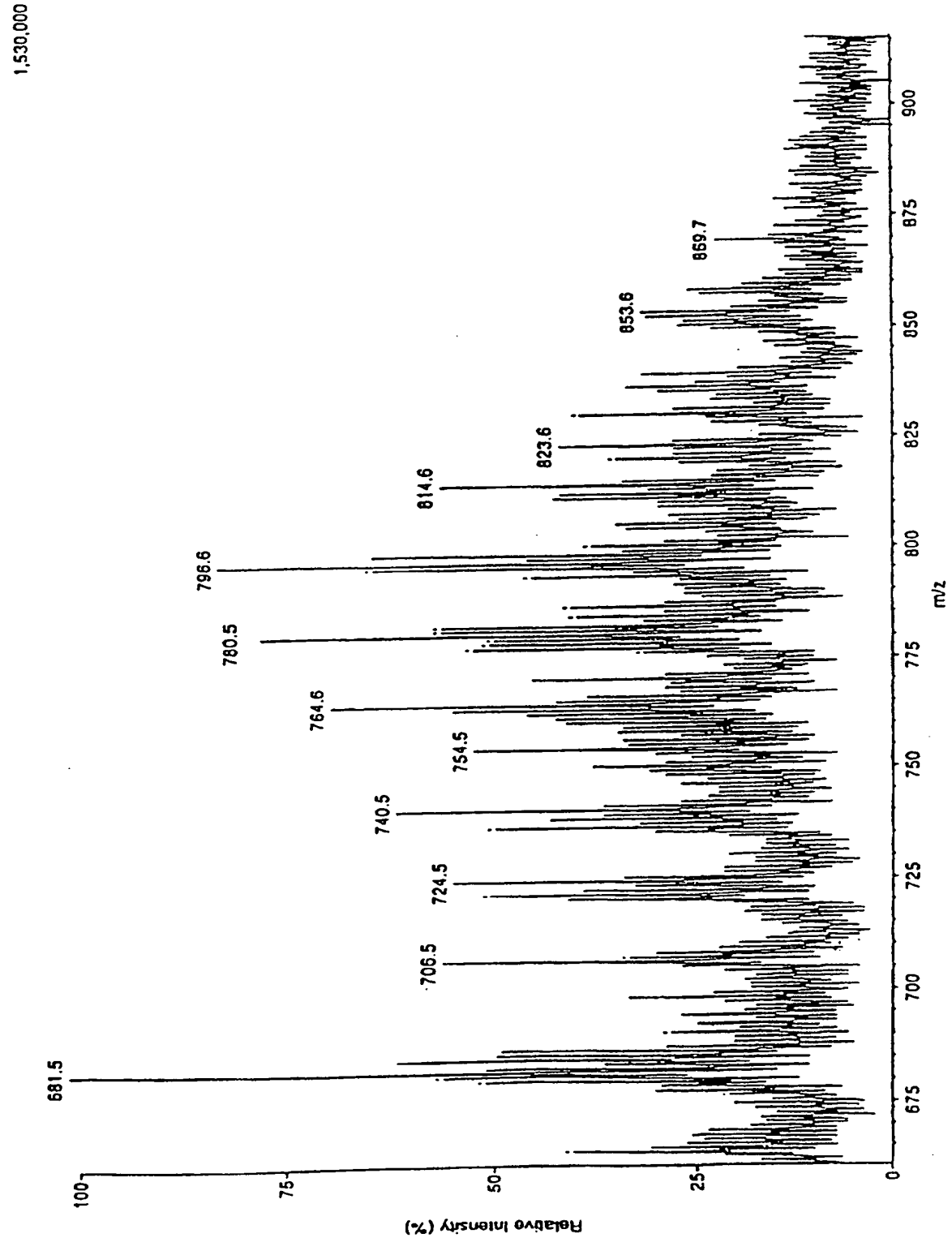


Figure 5

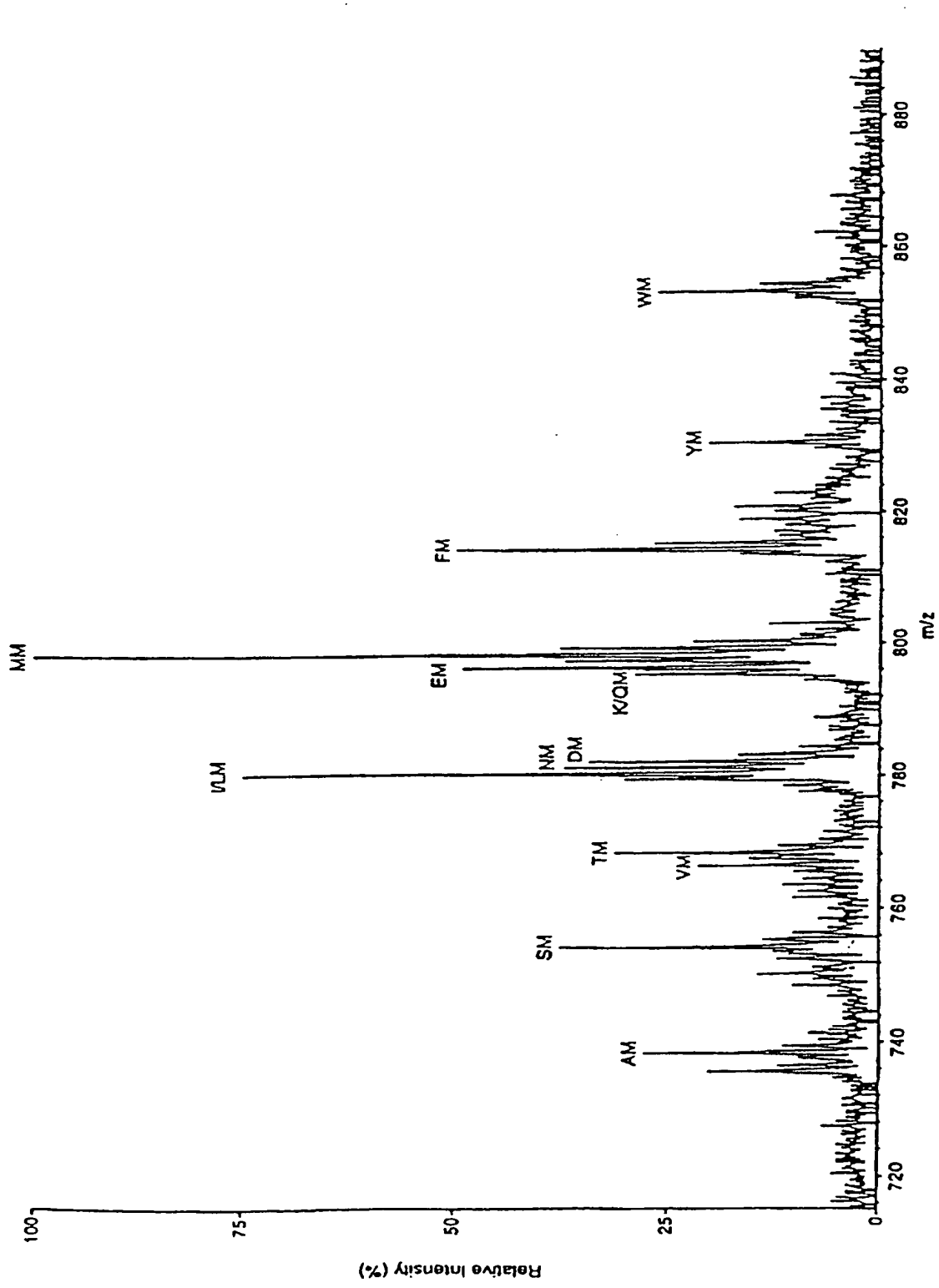
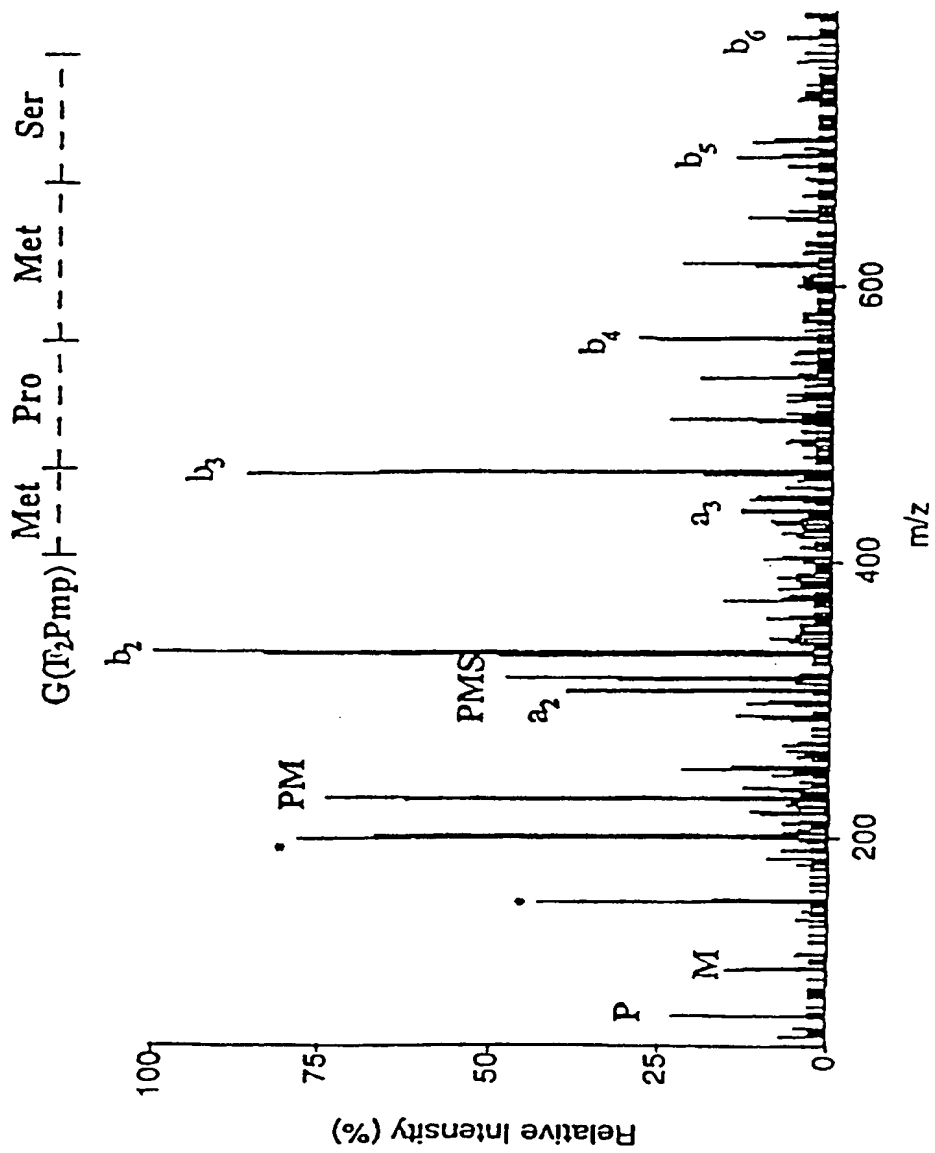


Figure 6



Apendix A1

The Mass Peak Identification Software (MPIS) provides a framework for using mass peak collections for identification of compounds in a combinatorial libraries. It is an Excel macro based procedure which identifies unknown mass peaks by automatically mapping them against a known library of mass peaks. Parameters for threshold filtering and fit variance are adjustable by the scientist for precision peak identification. Each peak found the standard library peak is presented as well as the closeness of fit between the unknown and library peaks. Known peak libraries are stored as PC files; so that, any desired library can be conveniently submitted for this procedure.

Apendix A2

```
' Mass Peak Indentification Software (MPIS)
' Masspec Macro
' Macro written 04/15/96 by Chris Goeller, Jake Raab
' Excel Version 5.0 Macro Visual Basic
'
Sub Masspec()
' Sort the data
    Columns("A:D").Select
    Selection.Sort Key1:=Range("A2"), Order1:=xlAscending, Header:= _
        xlGuess, OrderCustom:=1, MatchCase:=False, Orientation:= _
        xlTopToBottom
' Set up the headers for the new columns
    Columns("D:D").Select
    Selection.Insert Shift:=xlToRight
    Range("D1").Select
    ActiveCell.FormulaR1C1 = "Present"
    Range("E1").Select
    Selection.EntireColumn.Insert
    Range("E1").Select
    ActiveCell.FormulaR1C1 = "Variance"
    Range("F1").Select
    Selection.EntireColumn.Insert
    Range("F1").Select
    ActiveCell.FormulaR1C1 = "New"
    Range("G1").Select
    Selection.EntireColumn.Insert
    ActiveCell.FormulaR1C1 = "Not present"
    Columns("G:G").EntireColumn.AutoFit
'     ActiveCell.Offset(-6, -1).Range("A1").Select
     ActiveSheet.Cells(2, 1).Select

nextt:
'    ActiveCell.Offset(1, -1).Range("A1").Select
    ActiveCell.Offset(1, 0).Range("A1").Select

firsst:
    UName = ActiveCell.FormulaR1C1
    If UName = "" Then GoTo part2

' Get the mass spec value in first column
    num = UName

' Get the value in second column (*)
    ActiveCell.Offset(0, 1).Range("A1").Select
    t$ = ActiveCell.FormulaR1C1
    i = 0
    If t$ = "*" Then
            i = 1
    End If
' move back to first column
            ActiveCell.Offset(0, -1).Range("A1").Select
    If i = 0 Then GoTo nextt
' get mass above it
            ActiveCell.Offset(-1, 1).Range("A1").Select
            ta$ = ActiveCell.FormulaR1C1
            If ta$ = "*" Then
                ActiveCell.Offset(1, -1).Range("A1").Select
                idifab = 99
                GoTo below
            End If
' check if not library and already assigned a peak
            ActiveCell.Offset(0, 2).Range("A1").Select
```

Apendix A3

```
                ta$ = ActiveCell.FormulaR1C1

                If ta$ <> "" Then
                    ActiveCell.Offset(1, -3).Range("A1").Select
                    idifab = 99
                    GoTo below
                End If
                 ActiveCell.Offset(0, -1).Range("A1").Select
                ActiveCell.Offset(0, -3).Range("A1").Select
                numab = ActiveCell.FormulaR1C1
                idifab = num - numab
                ActiveCell.Offset(1, 0).Range("A1").Select
' get mass below it
below:
                ActiveCell.Offset(1, 1).Range("A1").Select
                ta$ = ActiveCell.FormulaR1C1
                If ta$ = "*" Then
                    ActiveCell.Offset(-1, -1).Range("A1").Select
                    idifbe = 99
                    GoTo compare
                End If
                ActiveCell.Offset(0, -1).Range("A1").Select
                chkit$ = ActiveCell.FormulaR1C1
                If chkit$ = "" Then
                      ActiveCell.Offset(-1, 0).Range("A1").Select
                    idifbe = 99
                    GoTo compare
                    End If
                numbe = ActiveCell.FormulaR1C1
                idifbe = numbe - num
                ActiveCell.Offset(-1, 0).Range("A1").Select
compare:
        If idifab = 99 And idifbe = 99 Then GoTo nextt
        If idifab < 1 And idifbe > 1 Then
           Selection.Cut
           ActiveCell.Offset(-1, 3).Range("A1").Select
           ActiveSheet.Paste
           ActiveCell.Offset(1, -3).Range("A1:D1").Select
            Selection.Delete Shift:=xlUp
           ActiveCell.Offset(-1, 0).Range("A1").Select
        End If
         If idifab > 1 And idifbe < 1 Then
           Selection.Cut
           ActiveCell.Offset(1, 3).Range("A1").Select
           ActiveSheet.Paste
           ActiveCell.Offset(-1, -3).Range("A1:D1").Select
            Selection.Delete Shift:=xlUp
                      ActiveCell.Offset(-1, 0).Range("A1").Select
        End If
            If idifab < 1 And idifbe < 1 And idifab < idifbe Then
           Selection.Cut
           ActiveCell.Offset(-1, 3).Range("A1").Select
           ActiveSheet.Paste
           ActiveCell.Offset(1, -3).Range("A1:D1").Select
            Selection.Delete Shift:=xlUp
           ActiveCell.Offset(-1, 0).Range("A1").Select
        End If
         If idifab < 1 And idifbe < 1 And idifbe <= idifab Then
           Selection.Cut
           ActiveCell.Offset(1, 3).Range("A1").Select
           ActiveSheet.Paste
           ActiveCell.Offset(-1, -3).Range("A1:D1").Select
            Selection.Delete Shift:=xlUp
```

Apendix A4

```
                        ActiveCell.Offset(-1, 0).Range("A1").Select
        End If
    GoTo nextt
' Part 2 fills in the new peak column with *
part2:
     ActiveSheet.Cells(2, 1).Select
nextt2:
    ActiveCell.Offset(1, 0).Range("A1").Select
    UName = ActiveCell.FormulaR1C1
    If UName = "" Then GoTo part2a
' Get the value in second column (*)
     ActiveCell.Offset(0, 1).Range("A1").Select
    t$ = ActiveCell.FormulaR1C1
    i = 0
    If t$ = "*" Then
         ActiveCell.Offset(0, -1).Range("A1").Select
            GoTo nextt2
    End If
     ActiveCell.Offset(0, 2).Range("A1").Select
     n$ = ActiveCell.FormulaR1C1
     If n$ <> "" Then
        ActiveCell.Offset(0, -3).Range("A1").Select
        GoTo nextt2
     End If
        ActiveCell.Offset(0, 2).Range("A1").Select
        ActiveCell.FormulaR1C1 = "*"
        ActiveCell.Offset(0, -5).Range("A1").Select
        GoTo nextt2
' Part 2a fills in the variance column with *
part2a:
     ActiveSheet.Cells(2, 1).Select
nextt2a:
    ActiveCell.Offset(1, 0).Range("A1").Select
    UName = ActiveCell.FormulaR1C1
    If UName = "" Then GoTo part3
    num = UName
' Get the value in second column (*)
     ActiveCell.Offset(0, 1).Range("A1").Select
    t$ = ActiveCell.FormulaR1C1
    i = 0
    If t$ = "*" Then
         ActiveCell.Offset(0, -1).Range("A1").Select
            GoTo nextt2a
    End If
     ActiveCell.Offset(0, 2).Range("A1").Select
     n$ = ActiveCell.FormulaR1C1
     If n$ = "" Then
        ActiveCell.Offset(0, -3).Range("A1").Select
        GoTo nextt2a
     End If
     num2 = ActiveCell.FormulaR1C1
     Var = num2 - num
        ActiveCell.Offset(0, 1).Range("A1").Select
        ActiveCell.FormulaR1C1 = Var
        ActiveCell.Offset(0, -4).Range("A1").Select
        GoTo nextt2a
' Part3 fills in the missing peaks columns in the library
part3:
     ActiveSheet.Cells(2, 1).Select
nextt3:
    ActiveCell.Offset(1, 0).Range("A1").Select
```

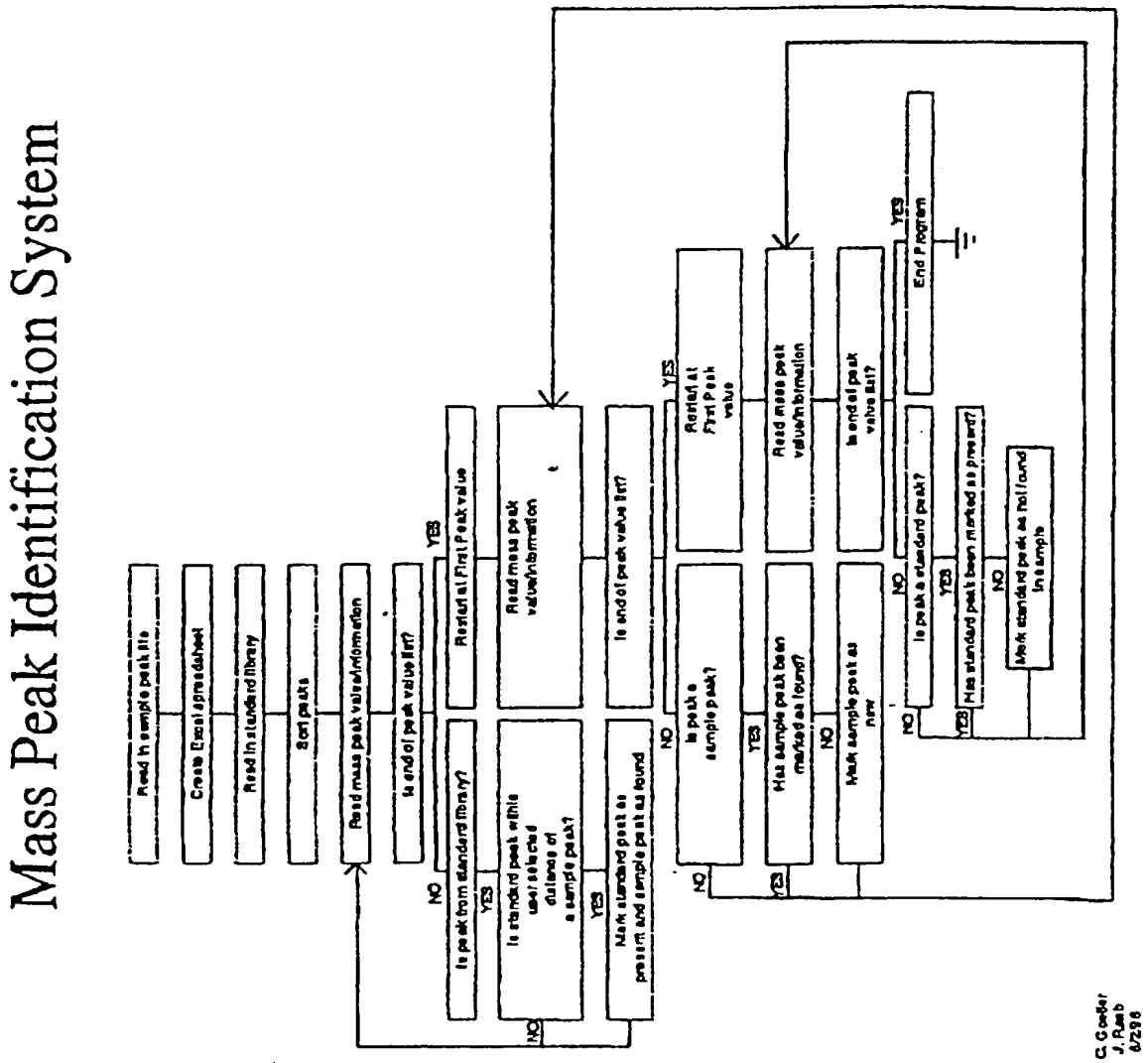
Apendix A5

```
    UName = ActiveCell.FormulaR1C1
    If UName = "" Then GoTo Endd
' Get the mass spec value in first column
    num = UName
' Get the value in second column (*)
    ActiveCell.Offset(0, 1).Range("A1").Select
    t$ = ActiveCell.FormulaR1C1
    i = 0
    If t$ = "*" Then
          i = 1
    End If
     If i = 1 Then
           ActiveCell.Offset(0, -1).Range("A1").Select
          Selection.Cut
          ActiveCell.Offset(0, 6).Range("A1").Select
          ActiveSheet.Paste
          ActiveCell.Offset(0, -5).Range("A1").Select
           ActiveCell.FormulaR1C1 = ""
           ActiveCell.Offset(0, -1).Range("A1").Select
           Else
            ActiveCell.Offset(0, -1).Range("A1").Select
  End If

   GoTo nextt3


Endd:
End Sub
```

Apendix B


Mass Peak Identification System
Read in sample peak file
Create Excel spreadsheet
Read in standard library
Sort peaks
Read mass peak value/information
Is end of peak value list?
YES
NO
Is peak from standard library?
YES
NO
Restart at First Peak value
Read mass peak value/information
Is standard peak within user selected distance of a sample peak?
YES
NO
Mark standard peak as present and sample peak as found
Is end of peak value list?
YES
NO
Restart at First Peak value
Is peak a sample peak?
YES
NO
Has sample peak been marked as found?
YES
NO
Read mass peak value/information
Mark sample peak as new
Is end of peak value list?
YES
NO
End Program
Is peak a standard peak?
YES
NO
Has standard peak been marked as present?
YES
NO
Mark standard peak as not found in sample
C. Gerker
J. Raab
6/96

Apendix C1

```
#!/bin/ksh
#
#           Shell Script for producing List of library
#           elements with increasing molecular weight.
#           ------------------------------------------
#
#           Copyright Ciba Corporation, Summit, NJ 07901
#
#   STEP 1
#
# Create a file where each line has the following format (tab separated):
#           <sequence number>
#           <combination id>
#           <MolWeight first AA>
#           <MolWeight second AA>
#           <Total MolWeight of library element>
#           <prime number code>
#
MolWeight
#
#   STEP 2
#
#   Sort the file of permutations according to prime number code
#   (colum 7).
#   Since all permutations of a given sequence will have
#   the same code output only one of them.
#
sort -k 6n -u <combinations >uniqcomb
#
#   STEP 3
#
#   Generate output list sorted according to mol weight
#
#           The list has the following format:
#           <mol weight> <number of lib elements> <AA codes>+
#
MkList
```

Apendix C2

```
#!/usr/local/bin/perl
#
#  PERL SCRIPTL: Generate elements of a combinatorial Library
#  with two changing amino acid positions.
#
#
#           Copyright Ciba Corporation, Summit, NJ 07901

#
#    Molecular Weight amino acids.
#
%Weight = (
                'G',  57.02,     # Glycine
                'A',  71.04,     # Alanine
                'S',  87.03,     # Serine
                'P',  97.05,                # Proline
                'V',  99.07,                # Valine
                'T', 101.05,                # Threonine
                'I', 113.08,                # Isoleucine
                'L', 113.08,                # Leucine
                'N', 114.04,                # Aspargine
                'D', 115.03,                # Aspartic acid
                'Q', 128.06,                # Glutamine
                'K', 128.09,                # Lysine
                'E', 129.04,                # Glutamic acid
                'M', 131.04,                # Methionine
                'H', 137.06,                # Histidine
                'F', 147.07,                # Phenylalanine
                'R', 156.10,                # Arginine
                'Y', 163.06,                # Tyrosine
                'W', 186.09,                # Tryptophan
                );
#
#   A prime number is assigned to each amino acid
#   Any combination of several AAs will have a unique product number, the
#   product of the prime numbers.
#
%Prime   = (
                'G',  2,          # Glycine
                'A',  3,        # Alanine
                'S',  5,        # Serine
                'P',  7,        # Proline
                'V', 11,        # Valine
                'T', 13,        # Threonine
                'I', 19,        # Isoleucine
                'L', 23,        # Leucine
                'N', 29,        # Aspargine
                'D', 31,        # Aspartic acid
                'Q', 37,        # Glutamine
                'K', 41,        # Lysine
                'E', 43,        # Glutamic acid
                'M', 47,        # Methionine
                'H', 53,        # Histidine
                'F', 59,        # Phenylalanine
                'R', 61,        # Arginine
                'Y', 67,        # Tyrosine
                'W', 71         # Tryptophan
                );

@NucAcid = keys %Weight;
open ( COMB, ">combinations" );
$iComb = 1;
foreach $first (@NucAcid) {
    foreach $second (@NucAcid) {
                $Comb = $first . $second;
                $Weight1 = $Weight{ $first  };
                $Weight2 = $Weight{ $second };
```

Apendix C3

```
                $Total = $Weight1 + $Weight2 + 551.18;
                $Code = $Prime( $first ) * $Prime ( $second );
                print COMB
              "$iComb\t$Comb\t$Weight1\t$Weight2\t$Total\t$Code\n" ;
                $iComb = $iComb + 1;
        };
    };
    close ( COMB );
```

Apendix C4

```
#!/usr/local/bin/perl
#
#  PERL SCRIPT to generate output list
#
#       Version 1.0  October 12, 1995  Mathic Thoma
#            Copyright Ciba Corporation, Summit, NJ 07901
#
#  List all Nucleic Acid cominations with the same molecular waight.
#  Order list according to molecular waight.
#
#  Input is read from file 'uniqcomb'
#  which is a list of unique combinations
#
open ( FILE, "uniqcomb" );
#
#  Read datafile and store info in assocciative array
#
while (<FILE>){
    ($LNum, $Seq, $w1, $w2, $Weight, $Code) = split ( "\t" );
    $SeqList{ $Weight } = $SeqList{ $Weight } . " " . $Seq;
    $nSeq{ $Weight} = $nSeq{ $Weight } + 1;
#    print "$total\t$weight{ $total }\n";
};
close ( FILE );
#
#  Sort the data according to
#
@MolWeight =  sort (keys %SeqList) ;
#
open (LIST, ">seqlist" );
foreach $wgt (@MolWeight) {
   printf LIST ("%7.2f: %3d  %s\n", $wgt, $nSeq{$wgt}, $SeqList{$wgt} )
};
close (LIST)
```

Apendix D1

# Program LIBANAL: Analysis of Combinatorial Libraries

## *Purpose*

In the context of LIBANAL a combinatorial library is the set of all compounds one can generate from a fixed substrate by inserting additional components at two different locations on the substrate. At each location any amino acid selected from a given list can be substituted. It is possible to substitute a version of the same amino acid at both locations. Each member of a library is identified by a two-letter code, the first letter is the one-letter abbreviation of the amino acid substituted at the first location, the second letter is the abbreviation of the amino acid substituted at the second location.

The purpose of the program is to determine how well the members of a library can be distinguished based on their molecular weight. To this end a table is generated listing in ascending order all molecular weights occurring among the members of the library. For each weight the number of members attaining this weight is given, as well as a list of their two-letter identifiers. If two members differ just in the order of the substituted amino acids they will have the same molecular weight. The table will list only one of the rearrangements.

## *Program Structure*

The program consists of three major steps:

1. create a list of all possible members by enumeration,
2. eliminate those members that differ from another member only by the arrangements of the substituted parts,
3. for each occurring molecular weight count the number of members with this weight and list their identifiers.

### Step 1

Create a list of all possible members by enumerating them. For each member carry along the following information:

1. a sequential number,
2. two letter identifier,
3. molecular weights of the substituted parts,
4. total molecular weight (substituted parts + substrate),
5. prime number code.

The prime number code of a member is computed by multiplying the prime numbers assigned to the substituted parts (the program assigns a different prime number to each of the possible substituting parts). The resulting number uniquely identifies the substituted parts, but not their sequence; i.e., the two substitutions AB and BA will result in the same code.

Apendix D2

### Step 2

From this list eliminate members that differ from another member only in the order of the substituted parts. This is achieved by sorting the list according to prime number code and retaining only the first element in a string of members with the same code.

### Step 3

A table is created with one row for each molecular weight occurring among the library members. This is achieved by sorting the output of the list generated in step 2 according to molecular weight. For each weight the number of members attaining this weight is counted, and the member identifiers are listed. Finally this table is printed.

**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 1/04 — A3

(11) International Publication Number: **WO 97/43301**

(43) International Publication Date: 20 November 1997 (20.11.97)

(21) International Application Number: PCT/EP97/02215

(22) International Filing Date: 30 April 1997 (30.04.97)

(30) Priority Data:
60/017,545 — 10 May 1996 (10.05.96) — US

(71) Applicant *(for all designated States except US)*: NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: WENNOGLE, Lawrence Paul [US/US]; 38 Inverrary Place, Annandale, NJ 08801 (US). KELLY, Michele, Ann [US/US]; 22 Van Ness Court, Maplewood, NJ 07040 (US). LIANG, Hongbin [CN/US]; Apartment A, 98 Gales Drive, New Providence, NJ 07974 (US). GOELLER, Christine [US/US]; 78 Westover Avenue, West Caldwell, NJ 07006 (US). THOMA, Hans, Mathis [CN/US]; 34 Ashland Road, Summit, NJ 07901 (US).

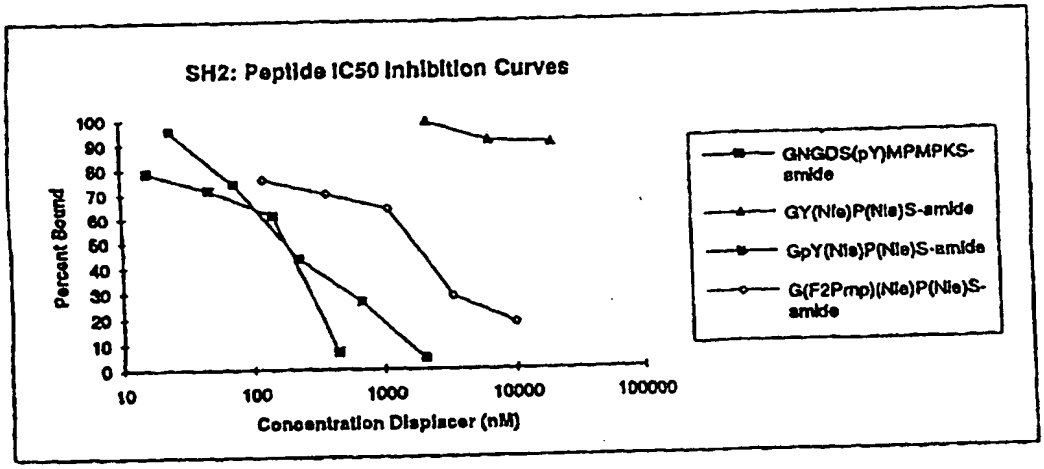
(74) Agent: ROTH, Bernhard, M.; Novartis AG, Patent- und Markenabteilung, Klybeckstrasse 141, CH-4002 Basel (CH).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(88) Date of publication of the international search report:
19 February 1998 (19.02.98)

(54) Title: IDENTIFICATION OF MEMBERS OF COMBINATORIAL LIBRARIES BY MASS SPECTROMETRY




| | IC50 (nM) Ave. | SD |
|---|---|---|
| GNGDS(pY)MPMPKS-amide | 188 | 53 |
| GY(Nle)P(Nle)S-amide | >20000 | |
| GpY(Nle)P(Nle)S-amide | 202 | 24 |
| G(F2Pmp)(Nle)P(Nle)S-amide | 2681 | 2958 |

(57) Abstract

The present invention is drawn to methods for characterizing the members of a combinatorial library which bind to a domain c interest. The method utilizes affinity selection in combination with mass spectrometry to provide rapid and efficient screening. The metho provides information on relative affinities and molecular weights of affinity-selected compounds. The methods find use in analyzing a types of combinatorial libraries.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | |
|---|---|
| AL | Albania |
| AM | Armenia |
| AT | Austria |
| AU | Australia |
| AZ | Azerbaijan |
| BA | Bosnia and Herzegovina |
| BB | Barbados |
| BE | Belgium |
| BF | Burkina Faso |
| BG | Bulgaria |
| BJ | Benin |
| BR | Brazil |
| BY | Belarus |
| CA | Canada |
| CF | Central African Republic |
| CG | Congo |
| CH | Switzerland |
| CI | Côte d'Ivoire |
| CM | Cameroon |
| CN | China |
| CU | Cuba |
| CZ | Czech Republic |
| DE | Germany |
| DK | Denmark |
| EE | Estonia |
| ES | Spain |
| FI | Finland |
| FR | France |
| GA | Gabon |
| GB | United Kingdom |
| GE | Georgia |
| GH | Ghana |
| GN | Guinea |
| GR | Greece |
| HU | Hungary |
| IE | Ireland |
| IL | Israel |
| IS | Iceland |
| IT | Italy |
| JP | Japan |
| KE | Kenya |
| KG | Kyrgyzstan |
| KP | Democratic People's Republic of Korea |
| KR | Republic of Korea |
| KZ | Kazakstan |
| LC | Saint Lucia |
| LI | Liechtenstein |
| LK | Sri Lanka |
| LR | Liberia |
| LS | Lesotho |
| LT | Lithuania |
| LU | Luxembourg |
| LV | Latvia |
| MC | Monaco |
| MD | Republic of Moldova |
| MG | Madagascar |
| MK | The former Yugoslav Republic of Macedonia |
| ML | Mali |
| MN | Mongolia |
| MR | Mauritania |
| MW | Malawi |
| MX | Mexico |
| NE | Niger |
| NL | Netherlands |
| NO | Norway |
| NZ | New Zealand |
| PL | Poland |
| PT | Portugal |
| RO | Romania |
| RU | Russian Federation |
| SD | Sudan |
| SE | Sweden |
| SG | Singapore |
| SI | Slovenia |
| SK | Slovakia |
| SN | Senegal |
| SZ | Swaziland |
| TD | Chad |
| TG | Togo |
| TJ | Tajikistan |
| TM | Turkmenistan |
| TR | Turkey |
| TT | Trinidad and Tobago |
| UA | Ukraine |
| UG | Uganda |
| US | United States of America |
| UZ | Uzbekistan |
| VN | Viet Nam |
| YU | Yugoslavia |
| ZW | Zimbabwe |

# INTERNATIONAL SEARCH REPORT

Inter ial Application No
PCT/EP 97/02215

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K1/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N G01N C12M C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 95 25737 A (PENN STATE RESEARCH FOUNDATION) 28 September 1995<br>see page 6, line 30 - page 20, line 25<br>see page 30, line 26 - page 31, line 26<br>--- | 1,2,7,8, 10,17 |
| X | JOURNAL OF THE AMERICAN CHEMICAL SOCIETY., vol. 117, 1995, DC US, pages 8859-8860, XP002043006<br>X. CHENG ET AL.: "USING ELECTROSPRAY IONIZATION FTICR MASS SPECTROMETRY TO STUDY COMPETITIVE BINDING OF INHIBITORS TO CARBONIC ANHYDRASE"<br>see the whole document, espcially page 8860, column 2<br>---<br>-/-- | 1,10,17, 18 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 9 October 1997 | 23. 12. 97 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | De Kok, A |

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Intern. al Application No
PCT/EP 97/02215

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 89, May 1992, WASHINGTON US, pages 4505-4509, XP002043007 R.N. ZUCKERMANN ET AL.: "IDENTIFICATION OF HIGHEST-AFFINITY LIGANDS BY AFFINITY SELECTION FROM EQUIMOLAR PEPTIDE MIXTURES GENERATED BY ROBOTIC SYNTHESIS" see page 4506, column 2 see page 4507 - page 4508 --- | 1,2,7, 11,14-16 |
| Y | ANGEWANDTE CHEMIE. INTERNATIONAL EDITION., vol. 31, no. 4, 1 April 1992, WEINHEIM DE, pages 367-383, XP000325208 G. JUNG ET AL.: "MULTIPLE PEPTIDE SYNTHESIS METHODS AND THEIR APPLICATIONS" see page 377, column 2 - page 380 | 1,19 |
| A | --- | 2,7,17, 18,21,25 |
| Y | CELL, vol. 72, no. 5, 12 March 1993, NA US, pages 767-778, XP000579645 Z SONGYANG ET AL.: "SH2 DOMAINS RECOGNIZE SPECIFIC PHOSPHOPEPTIDE SEQUENCES" see page 776, column 2 - page 777 | 1,19 |
| A | --- | 2-7,11, 12, 20-25, 27,28 |
| A | WO 95 18823 A (BETH ISRAEL HOSPITAL) 13 July 1995 see page 6, line 6 - line 25 see page 7, line 35 - line 38 see page 39, line 15 - line 20 --- | 1,3,4, 19,20 |
| A | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA., vol. 89, October 1992, WASHINGTON US, pages 9367-9371, XP002043008 R.J. SIMON ET AL.: "PEPTOIDS: A MODULAR APPROACH TO DRUG DESIGN" cited in the application --- | 8,11,17, 27,33 |
| P,X | WO 96 22530 A (CHIRON CORPORATION) 25 July 1996 see the whole document --- | 1-3,5,7, 10,11, 14,15,17 |
| P,X | GB 2 300 260 A (ZENECA LTD) 30 October 1996 see the whole document --- | 1,2,11, 14-16 |
| | -/-- | |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Inter. al Application No
PCT/EP 97/02215

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | BIOCHEMISTRY,<br>vol. 35, no. 36, 10 September 1996,<br>WASHINGTON US,<br>pages 11747-11755, XP002043009<br>M.A. KELLY ET AL.: "CHARACTERIZATION OF SH2-LIGAND INTERACTIONS VIA LIBRARY AFFINITY SELECTION WITH MASS SPECTROMETRIC DETECTION"<br>see the whole document<br>----- | 1-34 |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

national application No.
PCT/EP 97/02215

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. Claims 1-34
2. Claims 35-37
3. Claims 38-40

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-34

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/ EP 97/02215

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

1. Claims 1-34 A method for characterizing the members of a combinatorial library which bind to a domain of interest
2. Claims 35-37 A computer readable medium having a programm for the identification of compounds in a combinatorial library, the program comprising means for eliminating compounds that differ only in the order of substituted parts.
3. Claims 38-40 A computer readable medium having a program for the identification of compounds in a combinatorial library, the program comprising means for identifying unknown mass peaks by mapping known mass peaks.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern. al Application No
PCT/EP 97/02215

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9525737 A | 28-09-95 | EP 0751950 A | 08-01-97 |
| WO 9518823 A | 13-07-95 | US 5532167 A | 02-07-96 |
| WO 9622530 A | 25-07-96 | AU 4963796 A | 07-08-96 |
| | | CA 2209988 A | 25-07-96 |
| | | EP 0804729 A | 05-11-97 |
| GB 2300260 A | 30-10-96 | WO 9632642 A | 17-10-96 |